

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:18:38 ; Search time 9430.35 Seconds
(without alignments)
18036.975 Million cell updates/sec

Title: US-10-084-817-25

Perfect score: 5696

Sequence: 1 gactacgtgcactgcagg.....ggtttcttaaaaaaa 5696

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

BST:

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hct.*

9: gb_estci.*

10: gb_estd.*

11: gb_hct.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_nam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rtd.*

26: em_gss_dhg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3369.4	59.2	3371	29	AY398790 Homo sapi
2	2486.2	43.6	3371	29	AY398792 Mus muscu
3	2408.8	42.3	3273	29	AY398791 Pan trogl
4	1510.8	26.5	4109	11	AK080686 Mus muscu

C	5	1416.6	24.9	3530	11	AK089315
	6	1360.8	23.9	3412	29	AY405252
	7	1219	21.4	3240	29	AY405254
	8	854	15.0	939	13	BQ933301
	9	791.6	13.9	891	13	BQ925776
	10	790.2	13.9	896	13	BX390187
C	11	760	13.3	778	14	CA414667
	12	753	13.2	763	14	CK000113
	13	744.4	13.1	937	13	BQ934054
	14	736	12.9	904	13	BQ940133
C	15	735.6	12.9	750	13	BUG18192
C	16	726.8	12.8	918	13	BX328968
	17	725.6	12.8	3401	29	AY405253
	18	725.6	12.7	746	13	BX114562
C	19	723.6	12.7	756	12	BQ008487
C	20	717	12.6	757	14	CA449275
	21	714	12.5	794	12	BI758564
C	22	705.4	12.4	962	13	BX328967
	23	703.4	12.3	876	9	AU121101
C	24	694	12.2	712	14	CA420699
	25	691.6	12.1	721	12	BM718951
C	26	689.8	12.1	707	13	BU627598
	27	688.2	12.1	793	9	AI346330
C	28	686	12.0	701	14	CA446830
C	29	680.4	11.9	708	12	BM997135
	30	674.8	11.8	748	12	BG680063
C	31	668.2	11.7	714	14	CA428685
	32	667.8	11.7	752	14	CD511356
C	33	664.8	11.7	706	10	BF109743
C	34	662.8	11.6	718	9	AI453687
	35	650	11.4	660	9	AA149866
C	36	648.2	11.4	660	12	BM988626
C	37	641.6	11.3	693	10	BE673014
C	38	640.8	11.2	652	14	CA425104
C	39	640.4	11.2	660	14	CA431209
	40	640.2	11.2	651	12	BG939270
C	41	637.4	11.2	673	12	BQ002730
	42	636.6	11.2	672	14	CD512572
C	43	634.2	11.1	706	10	AW190828
	44	633.8	11.1	661	12	BM272400
C	45	630.8	11.1	655	12	BM971657

ALIGNMENTS

RESULT 1	AY398790	AY398790	3371 bp	DNA	linear	GSS 15-DEC-2003
LOCUS	AY398790	Homo sapiens THBS2 gene, VIRTUAL TRANSCRIPT, partial sequence,				
DEFINITION	AY398790	Genomic survey sequence.				
ACCESSION	AY398790	GI:39754782				
VERSION	AY398790.1	GSS.				
KEYWORDS						
SOURCE						
ORGANISM						
		Homo sapiens (human)				
REFERENCE						
AUTHORS						
		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
		Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
		Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL		Science 302 (5652), 1960-1963 (2003)				
PUBMED		14671302				
REFERENCE						
AUTHORS						
		Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
		Direct Submission				
TITLE						

JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA									
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.									
FEATURES	Location/Qualifiers									
source	1..3371									
gene	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>3371 /gene="THBS2" /locus_tag="HCM008"									
ORIGIN										
	Query Match	59.2%	Score 3369.4;	DB 29;	Length 3371;					
	Best Local Similarity	100.0%;	Pred. No. 0;							
	Matches 3370;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;	
Qy	132	ATGTCCTGGAGGCTGGTCTCTGCTCTGTGGGTGTGGCCAGCAGCAGCAAGCTGGTCAC	191							
Db	1	ATGTCCTGGAGGCTGGTCTCTGCTCTGTGGGTGTGGCCAGCAGCAGCAAGCTGGTCAC	60							
Qy	192	CAGGACAAAGACACGACCTTCGACCTTTTCAGTATCAGCAACATCAACCGCAAGACCACTT	251							
Db	61	CAGGACAAAGACACGACCTTCGACCTTTTCAGTATCAGCAACATCAACCGCAAGACCACTT	120							
Qy	252	GGCCCAAGCAGTTCCGGGGCCCGACCCCGCGCTGCCGGTTACCGCTTGGTGGCGTTT	311							
Db	121	GGCCCAAGCAGTTCCGGGGCCCGACCCCGCGCTGCCGGTTTACCGTTGGTGGCGTTT	180							
Qy	312	GACTTACATCCACCGGTGAACCGCAGATCACTTCAGCAAGATCACCAGATCATCGCGGAG	371							
Db	181	GACTTACATCCACCGGTGAACCGCAGATCACTTCAGCAAGATCACCAGATCATCGCGGAG	240							
Qy	372	AAGAGGGCTTCTCTCACCGCCCGAGTCAAGCAGGACGGCAAGTCCAGGGCACGGTG	431							
Db	241	AAGAGGGCTTCTCTCTCACCGCCCGAGTCAAGCAGGACGGCAAGTCCAGGGCACGGTG	300							
Qy	432	TTGGCTCTGGAGGGCCCCGGTCTCTCCAGAGGCAGTTTCGAGATCGTCTCCAACGGCCCC	491							
Db	301	TTGGCTCTGGAGGGCCCCGGTCTCTCCAGAGGCAGTTTCGAGATCGTCTCCAACGGCCCC	360							
Qy	492	CGCGACACGCTGGATCTCACTACTGGAATTGACGGCACCGGCATGTGGTCTCCCTGGAG	551							
Db	361	CGCGACACGCTGGATCTCACTACTGGAATTGACGGCACCGGCATGTGGTCTCCCTGGAG	420							
Qy	552	GACGTCCGCCCTGGCTGACTCCGAGTGGGAAGACGTCAACGTCAGGTGGCTGGCGAGACC	611							
Db	421	GACGTCCGCCCTGGCTGACTCCGAGTGGGAAGACGTCAACGTCAGGTGGCTGGCGAGACC	480							
Qy	612	TACAGCTTGCAAGTGGGCTGCACTCTATGACAGCTTCGCTCTGGACAGCCCTTCAC	671							
Db	481	TACAGCTTGCAAGTGGGCTGCACTCTATGACAGCTTCGCTCTGGACAGCCCTTCAC	540							
Qy	672	GAGCACTTCGAGCGGGAAGAGCGGATGTACGTGGCCAAAGGCTCTGCCAGAGAGAT	731							
Db	541	GAGCACTTCGAGCGGGAAGAGCGGATGTACGTGGCCAAAGGCTCTGCCAGAGAGAT	600							
Qy	732	CAC TTCAGGGTTTCTCTCAGAACCTCCACCTAGTGTGTTTGAATACTCTGTGAAGATAT	791							
Db	601	CAC TTCAGGGTTTCTCTCAGAACCTCCACCTAGTGTGTTTGAATACTCTGTGAAGATAT	560							
Qy	792	CTAAGCAAGAAGGGTTGCCAGCAAGGCCAGGAGCTGAGATCAACGCCATCATGTGAGAAC	851							
Db	661	CTAAGCAAGAAGGGTTGCCAGCAAGGCCAGGAGCTGAGATCAACGCCATCATGTGAGAAC	720							
Qy	852	ACAGAGACGCTCGCCTGGCTCGGATGTCACACCGAGTACGTGGGCCCCAGCTCAGAG	911							
Db	721	ACAGAGACGCTCGCCTGGCTCGGATGTCACACCGAGTACGTGGGCCCCAGCTCAGAG	780							
Qy	912	AGGAGGCCCGAGGTGTGCGAAACGCTCTGCCAGGAGCTCGGAAACATGTGCCAGGAGCTC	971							
Db	781	AGGAGGCCCGAGGTGTGCGAAACGCTCTGCCAGGAGCTCGGAAACATGTGCCAGGAGCTC	840							

2052 ACAGAAAGCAAGTGTGTGAGCCGCAAAACCCATGCAAGGACAGACACAACTGCCAC 2111
2121 ACAGAAAGCAAGTGTGTGAGCCGCAAAACCCATGCAAGGACAGACACAACTGCCAC 1980
2112 AAGCACGCGGAGTGATCTACCTGGGCGCACTTCAGCGACCCCATGATACAGTCCGAGTGC 2171
1981 AAGCACGCGGAGTGATCTACCTGGGCGCACTTCAGCGACCCCATGATACAGTCCGAGTGC 2040
2172 CAGACAGCTACGCGGGGCGACGGGCTCATCTCGGGGAGGACTCGGACCTGGACGGCTG 2231
2041 CAGACAGCTACGCGGGGCGACGGGCTCATCTCGGGGAGGACTCGGACCTGGACGGCTG 2100
2232 CCCAACCTCAATCTGGTCTGCGGCCCAACCAACGACCTACCTGCAATCAAGGATAACTGC 2291
2101 CCCAACCTCAATCTGGTCTGCGGCCCAACCAACGACCTACCTGCAATCAAGGATAACTGC 2160
2292 CCCCATCTGCCAATCTGGGCGAGGAAGACTTTGACAAGGACGGGATTTGGCGATGCTGT 2351
2161 CCCCATCTGCCAATCTGGGCGAGGAAGACTTTGACAAGGACGGGATTTGGCGATGCTGT 2220
2352 GATGATGACGATGACAAATGACCGTGTGACCGATGAGGAGGACAACTGCCAGCTCCTCTTC 2411
2221 GATGATGACGATGACAAATGACCGTGTGACCGATGAGGAGGACAACTGCCAGCTCCTCTTC 2280
2412 AATCCCGCCAGGCTGACTATGACAAAGGATGAGTTGGGACCGCTGTGACAACTGCCCT 2471
2281 AATCCCGCCAGGCTGACTATGACAAAGGATGAGTTGGGACCGCTGTGACAACTGCCCT 2340
2472 TACGTGACACACCTGCCAGATCGACACAGACAACAATGGAGAGGCTGACGCTGCTCC 2531
2341 TACGTGACACACCTGCCAGATCGACACAGACAACAATGGAGAGGCTGACGCTGCTCC 2400
2532 GTGACATTTGATGGGAGCGATGCTTCAATGAACGAGACAATTTGCCCTACGCTCAAC 2591
2401 GTGACATTTGATGGGAGCGATGCTTCAATGAACGAGACAATTTGCCCTACGCTCAAC 2460
2592 ACTGACAGAGGACACGATGGTGACGGTGTGGGGATCACTGTGACAACTGCCCTG 2651
2461 ACTGACAGAGGACACGATGGTGACGGTGTGGGGATCACTGTGACAACTGCCCTG 2520
2652 GTGACACACCTGTGACAGACCGACGTGGACAATGACCTTTGTTGGGACCAAGTGTGACAAC 2711
2521 GTGACACACCTGTGACAGACCGACGTGGACAATGACCTTTGTTGGGACCAAGTGTGACAAC 2580
2712 AACGAGGACATGATGACAGCGGCCACGAGACAACAGGACAACTGCCCTACATCTCC 2771
2581 AACGAGGACATGATGACAGCGGCCACGAGACAACAGGACAACTGCCCTACATCTCC 2640
2772 AACGCCAACCGCTGACCATGACAGAGACGGCCAGGCGACCGCTGTGACCTCTGATGAT 2831
2641 AACGCCAACCGCTGACCATGACAGAGACGGCCAGGCGACCGCTGTGACCTCTGATGAT 2700
2832 GACAAAGATGGGCTCCCGATGACAGGACAACTGCCGGCTTGTGTTCAACCCAGACGAG 2891
2701 GACAAAGATGGGCTCCCGATGACAGGACAACTGCCGGCTTGTGTTCAACCCAGACGAG 2760
2892 GAGGACTTGGACGGTGTGACAGCGGCTGATATTGTAAGATGATTTGACAATGACAAC 2951
2761 GAGGACTTGGACGGTGTGACAGCGGCTGATATTGTAAGATGATTTGACAATGACAAC 2820
2952 ATCCAGATTTGATGATGTGTCTGTAACAAATGCCATCATGAGGACAGACTTCAGG 3011
2821 ATCCAGATTTGATGATGTGTCTGTAACAAATGCCATCATGAGGACAGACTTCAGG 2880
3012 AACTTCCAGATGGTCCCTTGGATCCAAAGGACCAACCAATTTGATCCCACTGGGTC 3071
2881 AACTTCCAGATGGTCCCTTGGATCCAAAGGACCAACCAATTTGATCCCACTGGGTC 2940
3072 ATTGCCATCAAGGCAAGGAGCTGGTTTCAGACGCACTCGGACCCCGCATCGCTGTA 3131
2941 ATTGCCATCAAGGCAAGGAGCTGGTTTCAGACGCACTCGGACCCCGCATCGCTGTA 3000
3132 GGTTTTGACGAGTTGGGCTGTGGACTTCAGTGGGCACATTTCTACGTAACACTGACCGG 3191

Db 3001 GGTTTTGACGAGTTGGGCTGTGGACTTCAGTGGGCACATTTCTAGTAACACTGACCGG 3060
Qy 3192 GACACGACTATGCGGCTTCGTCTTTGGTTACCAAGTCAAGCAGCGCTTCTATGTGGTG 3251
Db 3061 GACACGACTATGCGGCTTCGTCTTTGGTTACCAAGTCAAGCAGCGCTTCTATGTGGTG 3120
Qy 3252 ATGTGGAGCAGGTGACCGCAGACCTACTGGGAGGACCCACCGGGCTTATGGCTAC 3311
Db 3121 ATGTGGAGCAGGTGACCGCAGACCTACTGGGAGGACCCACCGGGCTTATGGCTAC 3180
Qy 3312 TCCGCGGTGTCCTCAAGGTGGTGAACCTCACACCGGGACCGGCGAGCACTGAGGAAC 3371
Db 3181 TCCGCGGTGTCCTCAAGGTGGTGAACCTCACACCGGGACCGGCGAGCACTGAGGAAC 3240
Qy 3372 GCGTGTGGCACACGGGGAACACGCGGGGAGGTGCGAACCTTATGGACGACGCCCAGG 3431
Db 3241 GCGTGTGGCACACGGGGAACACGCGGGGAGGTGCGAACCTTATGGACGACGCCCAGG 3300
Qy 3432 AACATTGGCTGGAAGGACTACACGCGCTTATAGTGGCACCCTGACTCACAGGCCCAAGACT 3491
Db 3301 AACATTGGCTGGAAGGACTACACGCGCTTATAGTGGCACCCTGACTCACAGGCCCAAGACT 3360
Qy 3492 GGCTACATCAG 3502
Db 3361 GGCTACATCAG 3371

RESULT 2
AY398792 3371 bp DNA linear GSS 15-DEC-2003
LOCUS Mus musculus THB82 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY398792
VERSION AY398792.1 GI:39754784
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 3371)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3371)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
source location/Qualifiers
1..3371
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>3371
/gene="THB82"
/locus_tag="HCN0008"

gene
Query Match 43.6%; Score 2486.2; DB 29; Length 3371;
Best Local Similarity 83.6%; Pred. No. 0;
Matches 2818; Conservative 0; Mismatches 553; Indels 0; Gaps 0;

QY 132 ATGGTCTGGAGGCTGCTCTGCTGGCTCTCTGGGTGGCCAGCAGCAAGCTGGTCCAC 191
 Db 1 ATGCTCTGGGCACTGGCCCTGCTGCTGCTCTGGGCTATAGGCCCAAGAGCTTCTGCTGGTGC 60
 QY 192 CAGGCAAAAGACACAGACCTTTCGACCTTTTCAGTATACAGCAACATCAACCGCAAGCCATT 251
 Db 61 CACGTCAAGGACACTTCATTGACCTTTTCAGCATCAGCAACATTAACCGGAAGCCATC 120
 QY 252 GCGGCCAAGCAGTTCCGCGGCGCCGACCCCGCGTGGCGGTTCACCGCTTCGTGCGCTTT 311
 Db 121 GGTGCCAAGCAGTTCCGAGGCGCTGACCCCGGGTGGCCGCTACCGTTTGTACGGTTT 180
 QY 312 GACTCATCCCAACCGGTGAACAGAGATGACCTCAGCAAGATCACCAAGATCATGCGGCGAG 371
 Db 181 GACTATACCCCCAGTGAACACAGATGATCTCAACAGGATTGTCAAGCTTGCAGGAGA 240
 QY 372 AAGGAGGGCTTCTCTCAGCGGCCAGCTCAAGCAGGACGGCAAGTCCAGGGGACAGCTG 431
 Db 241 AAGGAGGGCTTCTCTCAGCGGCCAGCTCAAGCAGGACGGCAAGTCCAGGGGACAGCTC 300
 QY 432 TTGGCTCTGAGGGGCCGCTCTCTCCAGAGGAGTTCAGATCGTCTCCAAACGGCCCC 491
 Db 301 CTGGCTTGAAGGCCCGCCAGCTCCAGAGGAGTTCAGATCGTCTCCAAATGGCCCA 360
 QY 492 GCGGACACGTGATCTCACTCTGATTCAGCGCACCGCGCATGTGGTCTCCCTGGAG 551
 Db 361 GGGGACACTTTGACCTTCACTCTGGGTAGAGGCAATCAGCATACCACTTCTCTGGAG 420
 QY 552 GACGTCCGGCTGCTGACTCGCATGGAAGAACGTCAACCGTCAAGGTTGGTGGGAGACC 611
 Db 421 GATGTGGGCTGCTGACTCCAGTGAAGAAATGTGACTGTGCAAGTGGCCAGTGACACC 480
 QY 612 TACAGCTTGACGTGGCTGGACCTCATAGACAGCTTCGTCTGGACGAGCCCTTCTAC 671
 Db 481 TATAGCTGTATGGGCTGCGATCTTATCAGAGTGTCAACCTGGAAGAACCTTCTAT 540
 QY 672 GAGCACTTCAGCGGGAAGAGCGGATGTACGTGGCCAAAGGCTTCGCCAGAGAGAT 731
 Db 541 GAGCAGCTAGAAGTAGACAGGACGAGATGTACGTGGCCAAAGGTCATCTCGAGAGAT 600
 QY 732 CACTTCAGGGTTGCTTCAAGAGCTCCACTAGTGTGTAAGACTCTGTGGAGATATT 791
 Db 601 CACTTCAGGGGCTGTGTCAGAAATGTCCATCTCGTATTTGAGATCTCTGTGAAGATATC 660
 QY 792 CTAAGCAAGAGGTTTCCAGCAAGGCGCAGGAGCTGAGATCAACGCCATCAGTGAGAAC 851
 Db 661 TTAAGCAAGAAAGCTGTCAACACAGCAGGAGCTGAAGTCAACACCATCAGTGAACAT 720
 QY 852 ACAGAGCCTGCGCTGGGTGCGCATGTCAACACCGAGTACGTGGGCGCCAGCTCAGAG 911
 Db 721 ACAGAGACTCTCCATCTGAGCCCTTCAATCAACACAGATCTCGTGGTCCAGGGTGTGAG 780
 QY 912 AGGAGGCCGAGGTGTGCGAAACGCTCGTCCAGGAGCTGGGAAACATGGTCCAGGAGCTC 971
 Db 781 AAGGCACAGAGGTGTACGCACTCTCTCGAGGAGTTGAGCAACATGATGAATCAGCTC 840
 QY 972 TCGGGGCTCCAGTCTCTGTGAACACAGCTCAGCGAGAACCTCAGAGAGTGTGCAATGAT 1031
 Db 841 TCTGGAATGACGCTCATGTTGAAACAGCTGAGCAAGAACCTCGGAGAGTGTCTAGTGAT 900
 QY 1032 AACCAAGTTCTCTGGGAGCTCATTTGGTGGCCCTCTCAAGAACAGGAAATGTGAGCTTGC 1091
 Db 901 AACAGTTCTTTTGGAGCTCATTTGGGGGCCCTCTGAAGACAGAAACATGTACGCTGT 960
 QY 1092 TGGCAGGATCGCGGTTCTTTGGGAAATGAACGTTGGGTGGTGGACAGCTGCAACGAG 1151
 Db 961 GTGCAGAGGGGCCGAATCTTTGCAAGAAATGAACCTGGGTTGGATAGTTGTACCA 1020
 QY 1152 TGTACCTGCAAGAAATTTAAACCATTTGACCAAAATCACTGCGCCCGCTCAACCTGC 1211
 Db 1021 TGACCTTGCAAGAAATTTAAACCATGTGTCATCAGTCACTGCTACCTGCACTTGT 1080
 QY 1212 GCCAGTCCATCTTTTGGAAAGGCGAATGCTGCCCTTCTGCTCCATCTCGGTGACGGT 1271

Db 1081 GCCAACCCATCTTTTGGAAAGCGAGTGTGTCTCATCTCTGTTCACACTCTCGACACAGT 1140
 QY 1272 GAGGAGGCTGTGCTCTCGTGGGAGAGTGAACCGAGTCTCCGTGACGCTGTGGCTCTGG 1331
 Db 1141 GATGAGGCTGTGCTCTCGTGGGAGAGTGGACCGAGTGTCTGTCACTCTGTGGCTCTGG 1200
 QY 1332 ACCACGACAGAGAGCGGCTCTGTGAGCTCACACGAAACACTCTCTGGGCGCTTCCATC 1391
 Db 1201 ACCACGACAGAGAGCGGCTCTGTGATGTCAACAGCAACACTCTGGGCGCTTCCAT 1260
 QY 1392 CAGACAGGCTTGCAGTCTGAGCAAGTGTGACACCGCATCCCGCAGAGCGCGCTGG 1451
 Db 1261 CAGACAGGACATGCAAGCTTGGGCAATGTATACGAAATCCGTGAGATGAGAGCTGG 1320
 QY 1452 AGCACTGTGTCACTTTGCTTTCATGCTCTGTGACCTGTGAGTGTGGCAATATCACAGC 1511
 Db 1321 AGTCACTGTGTCACTTTGCTTTCATGCTCTCGTGTGATGTTGGAGTGTGCAATGTCA 1380
 QY 1512 ATCCGTCTCTGCACTCCCGAGTGGCCAGATGGGGGCAAGATTCGAAAGGAGTGGC 1571
 Db 1381 ATACGTCTCTGCAACTCACAGTGGCCCGAGATGGGTGGCAAGAACTGCAAGGGCAGCGC 1440
 QY 1572 CGGAGACCAAAAGCTTCCAGGGCGCCCATGATGCCCAATCGATGGCCGTGGAGCCCTGG 1631
 Db 1441 CGGAAACCAAAAGCTTGTGAGCGTGTATCCGTGCCAATGTGATGGCCGTGGAGCCCTGG 1500
 QY 1632 TCCCGTGTGCGCTGCACTGTCACTGTGCGGTGGATCCGGGAGCGCACCGGCTC 1691
 Db 1501 TCCCGTGTGTCAGCTGTGACAGTACCTGTGCTGGAGGATCCGTGAGCGCTCAGCTGT 1560
 QY 1692 TGCAACAGCCCTGAGCTCTAGTACGAGGGAAGGCTCGTGGGGGATGTGAGGAGCGT 1751
 Db 1561 TGCAACAGCCCTGAGCCCGAGTATGAGGGAAGGACTGTGTGCGGATGTGACAGAACAC 1620
 QY 1752 CAGATGTGCAACAGAGGAGTGGCCCGTGGATGGCTGTGTTATCCAAACCTCTCCG 1811
 Db 1621 CAAATGTCAACAGAGAGAGTCCCTTATGATGGGTGCTTATCCAAACCTGTTCCT 1680
 QY 1812 GAGGCCAGTGCAGCAGCTTCCCGATGGGTCTCTGTCATGCGGTCTCTGCGCTGTGGC 1871
 Db 1681 GAGGCCAGTGCAGCAGCTTCCCTGATGGGTCTCTGCTCTGTGGCTCTGCGCAGTGGC 1740
 QY 1872 TTCTTGGGCAATGGACCCACTGTGAGGACCTGACGAGTGTGCGCTGTGCTGGTCCCCGACATC 1931
 Db 1741 TTCTTGGGCAATGGTATACCCACTGTGAGGACCTGTGATGAGTGTGTGTGTCAGATAT 1800
 QY 1932 TGCTTCTCCACAGCAGGAGTGGCTCGCTGTGTCAACACTCAGCCTGTGGTTCACCTGCTG 1991
 Db 1801 TGCTTCTCACTAACAAAGCTCCCGCTGTGTCAACCAACCCCGGCTTCCACTGCTG 1860
 QY 1992 CCTGCGCGCCCGATACAGAGGAAACAGCCCTCGGGTGGCTCGGCTGGAAGACGCAAG 2051
 Db 1861 CCTTGTCCACACGCTACAGGGGAACCAACCTTGGGTGTGGCTGTGGAGTGTAGG 1920
 QY 2052 ACGGAAAGCAAGTGTGTGAGCCGAAAACCATGCAAGGACAGACACAACTGCCAC 2111
 Db 1921 ACAGAAACCAATGTGTGAGCCAGAGATCCATGTAAGGACAGACTCACAGCTGCCAC 1980
 QY 2112 AAGCACGCGAGTGCATCTTACTTGGGCACTTTCAGGAGCCCATGTACAGTGGAGTGC 2171
 Db 1981 AAGAAATGCAGATGTGATCTACTTGGGCACTTGTAGTGAACCCCATGTACAAAGTGTGAGTGC 2040
 QY 2172 CAGACAGGCTACCGGGCGAGCGGCTCATCTGCGGGGAGGACTCGGACTGGAAGCGGCTGG 2231
 Db 2041 CAGATTGGCTACGAGGTGTGAGGCTCATCTGCGGGGAGGACTCAGACTGGATGGCTGG 2100
 QY 2232 CCCAAGCTCAATCTGTGCTGGCCACCAACGCGCACTTACCTACCTGATCAAGGATTAAGTGC 2291
 Db 2101 CCCAAGCAAACTGTGTGTCTACTAATGCCCACTTACCTACCTGATCAAGGACAACTGC 2160
 QY 2292 CCCCATCTGCCAATCTTGGGCGAGGAGACTTTCAGAGGAGCGGATTCGCGATGCTGT 2351

2161	CCCAAAGCTGCCAAATTC	CGGCGAGGAGGATTTT	TGATAAGGATCGGAATCGGAGATGCTTGT	2220
2352	GATGATGACGATGACAAATG	ACCGTGTGACCGATGAGAAGGACAACTGCCAGTCTCTCTTC	2411	
2221	GACGAGGACGATGACAAATG	ACCGTGTGAGCGATGAGAAGGACAAATTCGCCAGTCTCTCTTC	2280	
2412	AATCCCGCCGAGCTGAC	TATGACAAGGATGAGGTTGGGAGCCGCTGTGACAACTGCCCT	2471	
2281	AATCCCGCTCAATTAG	ACTATGACAAGGATGAGGTTGGAGACCGCTGTGACAACTGCCCC	2340	
2472	TACGTGCAAAACCCCTG	CCAGATCGACACAGACAACAATGAGAGAGGATGACCGCTGCTCC	2531	
2341	TATGTGCACAACCCAG	CACAGATCGACACAGACAACAATGCGCAGGAGGATGCGCTGCTCT	2400	
2532	GTGGCAATTGATGGGAC	AGATGCTTCAATGAACGAGACAATTTGTCCTACGCTCTACAAC	2591	
2401	GTGGCAATTGACGGAG	ACGATGTTTCAATGAGCGGAGACAATTTGTCCTATGCTACAAAC	2460	
2592	ACTGACAGAGGACACG	GAATGATGCGGTGGGGATCACTGTGACAACTGCCCCCTG	2655	
2461	ACTGACAGAGAGATAC	GGATGATGCGGTGGGATGACCACTGTGACAAATTTGCTCTCTG	2520	
2652	GTGCACAAACCCCTG	ACACGACCGGACCAATGACCTTGTGGGACCACTGTGACAAAC	2711	
2521	ATGCACAAACCCAG	ATCAGATCGATCAGGACAATGATCTCGTTGGAGACCACTGTGACAAAC	2580	
2712	AACGAGACATAGATG	ACGACGCGCCACAGAACAAACGAGGACAACTGCCCTCATCTCTCC	2771	
2581	AATGAGGACATAGAT	GATGATGATCGGCCACCAAGACAACCAACGACAACTGCCCATACATCTCC	2840	
2772	AACGCCAAACAGGCT	GACCATGATGACAGACCGGCGACGGCGACCGCTGTGACCCCTGATGAT	2831	
2641	AACTCCAAACAGGCT	GACCATGATGACACGCGCAAGGGCGATGCTCGCTGCGACTCTGATGAT	2700	
2832	GACAAAGATGCGCT	CCCCGATGACAGGACAACCTGCCCGGTGTGTGTTCAAACCGACACCAG	2891	
2701	GACAATGATGGTGT	TCAGATGACAGGACAACCTGTGGGTGTGTGTTCAAACCGACACCAG	2760	
2892	GAGGACTTGACCGGT	GATGAGACGGGTGATTTGTTAAAGATGATTTTGACAATGACAAAC	2951	
2761	GAAGACTCGGACGT	GACGCGCGAGGTGACATTTGTAAAGATGACTTTGACAATGATAT	2820	
2952	ATCCAGATATTGATG	ATGATGATGTCCTGAAAAAACAATGCCATCAGTGAGACAGACTTCAGG	3011	
2821	GTCCAGATATTGATG	ATGATGATGTGTCCTTGAGAACAAATGCCATCACTGAGACAGACTTCAGA	2880	
3012	AACTTCAGATGTGCC	CTTGGATCCCAAAAGGACACCAACCAAAATGATCCCAACTGGGTC	3071	
2881	AACTTCAGATGTGCC	CTTGGATCCCAAGGGACCAACAATAATGATCCCAACTGGGA	2940	
3072	ATTGCGCATCAAGG	CAAGGAGCTGGTTACAGACGCACTCGGACCCCGGCATCGCTGTA	3131	
2941	ATTGCTCACCAGG	CAAGGAGCTGGTGACAGACGACAACTCAGACCCCTGGCATCGCTGTA	3000	
3132	GGTTTTGACAGATT	TGGGCTGTGGACTTCAGTGGGCACTTCTACGTAAACACTGACCCGG	3191	
3001	GGTTTTGACAGATT	TGGGCTGTGGACTTCAGTGGGCACTTCTATGTCAACACTGACCCGG	3060	
3192	GACGACGACTATG	CGCGCTTCGTTTGGTTACAGTCAAGACGCGCTTCTATGTGGTG	3251	
3061	GATGATGACTACG	CTGCTTGTCTTTGGCTATCAGTCAAGACGCGCTTCTATGTGGTG	3120	
3252	ATGTGAAGCAGGTG	ACGACAGACTACTGGGAGGACCAAGCCCAACCGCGGCTATGGCTAC	3311	
3121	ATGTGAAGCAGGTG	ACGACAGACTACTGGGAGGACCAAGCCCAACCGGCTTACCGGCTAC	3180	
3312	TCGGCGGTGTCCT	CAGGTGGTGAACCTCCACCAACGGGACGGCGGACCTTGAGGAAC	3371	
3181	TCGTGTGTCACTCA	AAGTGGTAACTCCACGACTGGTACTGGCGAGACCTTGAGGAAT	3240	
3372	GCCTGTGGCACAC	GGGGAACAGCCCGGGCAGGTGCGAAACCTTATGGCACGCCACCGG	3431	
3241	GCCTGTGGCACAC	GGGGAACACAGAAAGGCAGGTCCGGACTCTATGGCATGACCCCAAA	3300	

Qy	3432	AACATTTGGCTGAAGGACTACACGGCTATAGTGGGACCTGACTCAGAGGGCCAAAGACT	3439
Db	3301	AACATTTGGCTGAAGGACTACACTGCTACAGTGGGACCTGATTCACAGGGCTAAGACA	3360
Qy	3492	GGCTACATCAG	3502
Db	3361	GGCTACATGAG	3371
RESULT 3	AY398791	3273 bp	DNA
LOCUS	Pan troglodytes THBS2 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.		linear GSS 15-DEC-2003
DEFINITION	AY398791		
ACCESSION	AY398791.1	GI:39754783	
VERSION			
KEYWORDS	GSS.		
SOURCE	Pan troglodytes (chimpanzee)		
ORGANISM	Pan troglodytes		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.		
AUTHORS	1. (bases 1 to 3273)		
	Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejaraiwal,A., Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B., Fieriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios		
JOURNAL	Science 302 (5652), 1960-1963 (2003)		
PUBMED	14671302		
REFERENCE	2. (bases 1 to 3273)		
AUTHORS	Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejaraiwal,A., Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B., Fieriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.		
FEATURES	Location/Qualifiers		
source	1..3273		
gene	/organism="Pan troglodytes"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9598"		
	<1..>3273		
	/gene="THBS2"		
	/locus_tag="HCM0008"		
ORIGIN			
Query Match	42.3%;	Score 2408.8;	DB 29; Length 3273;
Best Local Similarity	73.8%;	Pred. No. 0;	
Matches 2416;	Conservative 0;	Mismatches 857;	Indels 0; Gaps 0;
Qy	132	ATGGTCTGGAGCTGGTCTCTGCTGGCTCTGTGGGTGTGGCCAGCACGCAAGCTGCTCAC	191
Db	1	ATGGTCTGGAGCTGGTCTCTGCTGGCTCTGTGGGTGTGGCCAGCACGCAAGNNNNNNN	60
Qy	192	CAGGACAAAGACACGACCTTCGACCTTTTCAGTATCAGCAACATCAACCGCAAGACCAT	251
Db	61	NN	120
Qy	252	GGCGCAAGCAGTTCGCGGGCCGACCCCGCGTTCACCGCTTCGTCGCGCTTT	311
Db	121	NN	180
Qy	312	GACTACATCCACCGGTGAACGACGATGACCTCAGCAAGATCAACAAGATCATGGGCAG	371
Db	181	NN	240
Qy	372	AAGGAGGGCTTCTTCCTCAGCGGCCAGCTCAAGCAGGACGCGCAAGTCAGGGGCGCGCTG	431


```

/tissue_type="cortex"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
217..3733
misc_feature
  note="THROMBOSPONDIN 1 PRECURSOR homolog [Mus musculus]
  (SWISSPROT|P35441, evidence: FASTV, 99.7%ID, 100%length,
  match=3514)
  putative"

```

ORIGIN

```

Query Match      26.5%; Score 1510.8; DB 11; Length 4109;
Best Local Similarity 66.4%; Pred. No. 4.5e-216;
Matches 2258; Conservative 0; Mismatches 1112; Indels 30; Gaps 5;

QY 270 GGGCCGACCCCGGCTTACCGCTTCCGCTTGGCGCTTGGACTACATCCACCCGCTG 329
DB 367 GGCAGATCTATCCAGCCGCTTCCGATTTAGAAATGCCAACCTGATCCCCGCTGTG 426
QY 330 AACGAGATGACCTCAGCAAGATCAACAAGATATCGCGCAGAGAGGGCTTTCTCTC 389
DB 427 CCGGATGACAAGTTTCAAGACCTACTGGACGCTGTGTGGCCGACAAAGGCTTCATCTC 486
QY 390 AGGGCCACCTCAAGCAG-GACGGCAAGTCCAGGGCACGCTGTGGCTCTGGAGGGCC 448
DB 487 CTGGCTTCTTGGAGCAGATGAAAAAACCAGGGGCACACTCTCTGCTGTGGAACGGAA 546
QY 449 CGGTCTCTCCAGAGCAGTTCCAGATCGTCTCAACGGCCCGCGACACGCTGGATCT 508
DB 547 AGAAAAACCTGGCCAGATCTTCACTAGTGTGGTCTCAACGGCAAGCTGGCACCTTGGACCT 606
QY 509 CACCTACTGGATGAGCGCACCGGATGTGCTCTCTGGAGACGCTGCGCCCTGGCTGA 568
DB 607 GAGCTGAGCTTCCAGAGGAGCAACAGTGTGTCTAGTGGAGGAGCTCTCTTGGCCAC 666
QY 569 CTCGAGTGGAGAACCTCACCCTGCAAGTGGTGGCGAGACCTACAGCTTGCACGTGG 628
DB 667 TGGCAGTGGAGAGCATACGCTGTGTGTTCAAGAGGACCGGGCTCAACTCTACATAGA 726
QY 629 CTGCGACCTCATAGACAGCTTCCCTCTGGACGAGCCCTCTACGACACCTGCGAGGGGA 688
DB 727 CTGTGATAGATGGAGCGGAGCTGGATGTATCCATCA-GAGCATCTTACCAGGG 785
QY 689 AAAGACCGGATGTACGTGGCCAAAGCTCTGCGCAGAGAGTCTC-----ACTTCAGGG 741
DB 786 ATCTGCGCAGCTTGCAGGCTCCGAGTTGCAAGGAGAGATGTAATGACAAATTTTCAGG 845
QY 742 GTTGTCTTCCAGACGCTCCACTAGTGTGTAAGTCTGTGGAGATATTTTAAAGCAGA 801
DB 846 GGGTGTCTGAGAAATGAGGTTGTCTTTGGAACACCCCAAGAGACATTTCTCAGGAACA 905
QY 802 AGGGTTGCCAGAACGCGCAGGAGCTGAGATCAACGCCATCAGTGAGAACACAGAGACGC 861
DB 906 AAGGCTGC---TCCAGCTCAGCTACCAACGCTCTTTTACCTTTGACAAACGCTGGTGA 962
QY 862 TGGCCTGGTCTCGCATGTCCACACGAGTACGTGGGCGCCCGAGCTCAGAGAGGAGGCCG 921
DB 963 AGGTTCCAGCCCTGTATCCGACCAACTATATCGGCCCAAAACGAGACCTTCAAG 1022
QY 922 AGGTGTGCGAACGCTCGTGGAGAGCTGGGAAACATGGTCCAGAGCTCTCGGGGCTCC 981
DB 1023 CTATCTGTGGCTCTCTCTGTGATGAATCTATCCAGCAGTGGTCTTGGAACTGAAGGCGCTGC 1082
QY 982 AGTCTCTGTGAACAGCTCAGCGAGAACCTCAAGAGAGTGTGCAATGATTAACAGATTTC 1041
DB 1083 GCACCATCTGACCACTCTCGAGGACAGCATCCGAAAGTACGAGAGAGAACAGAGAGC 1142
QY 1042 TCTGGAGGCTCATTTGTTGGCCCTCTTAAGACAGAAACATGTCACTGTCTGGCAGGATG 1101
DB 1143 TGTCTAGTGAAGTGAAGCGGCTTCCCTC-----TGTCTTCACAATG 1184
QY 1102 GCGGTTCTTTGGGAAATGAAACGTTGGTGTGGACAGCTGCACACGCTGTACTCTCA 1161
DB 1185 GAGTCCAGTACAAAGAACACAGGAGTGGACTGTAGACAGTTTGACAGAGTGTCACTGCC 1244

```

```

QY 1162 AGAAATTTAAACCAATTGGCCACCAATACCTGCCCGCCCTGCAACCTGCCAGTCCAT 1221
DB 1245 AGAAGCTCGGTTACCATCTGCAAAAAGGTGCTCTCCCATCATGCCCTGTCTCAACGCA 1304
QY 1222 CTTTGTGGAAGCGCAATGCTGCCCTTCTCTGCCCTCCACTCGGTGGAGGAGGAGGCT 1281
DB 1305 CAGTTCTCTGATGGTGAATGCTGCCACGGTGTGGCCCGCAGGACTCTGCTCAGCATGCT 1364
QY 1282 GTCTCCGTGGGAGAGTGAGACCCAGTGTCTCGTGAGCTGTGGCTTGGGACCCAGCAGA 1341
DB 1365 GGTCTCCCTGGTCTGAGTGGACCTCTGCTCTGCCCATGTGGCAATGGAATTCAGCAAC 1424
QY 1342 GAGGCGCGTCTGTGACGTCAACAGCAACACTGCTTTGGGGCCCTCCATCCAGACACGGG 1401
DB 1425 GTGGTCTGTTCTGTGACAGCTCAACAACAGATGCGAGGGCTCTTCGTACAGACAAGGA 1484
QY 1402 CTTGAGTCTGAGCAAGTGTGACACCCGATCCGGAGGACGGCGGTGGAGCCACTGGT 1461
DB 1485 CTTGCCACATTCAGGAGTGTGACAAAAGATTTAAACAGGATGGTGGCTGGAGTCACTGGT 1544
QY 1462 CACCTTGGTCTTCTGATCTGTGACCTGTGGAGTTGGCAATATCACACGCTATCCGCTCT 1521
DB 1545 CTCCATGGTCTGCTCTGTTCTGTGACCTGTGAGCTGTGATCAAGGATTCGGCTCT 1604
QY 1522 GCAACTCCCCAGTGTCCCGAGATGGGGGCAAGATTCGAAAGGAGTGGCCGGGAGACCA 1581
DB 1605 GCAACTCCCCAGTGTCCCGAGATGAAACGGGAAGCCCTGTGAAGGTGAAGCCCGGGAGACA 1664
QY 1582 AAGCCTGCCAGGCGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1641
DB 1665 AAGCCTGCCAGGAGACGCTGCTCCCAATTAATGAGAGCTGGGGTCCCTGGTCAACATGGG 1724
QY 1642 CGGCTGTGATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1701
DB 1725 ACATCTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1784
QY 1702 CTGAGCTCAGTACGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1761
DB 1785 CCACACCCAGTTTGGAGGCAAGACATGCTGTGTCGATGTGACAGAAAATCAAGTTTGA 1844
QY 1762 ACAAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1821
DB 1845 ACAAGCAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1904
QY 1822 GCAGCAGCTTCCCGGATGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1881
DB 1905 GTACTAGCTACCTGATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1964
QY 1882 ATGGACCCACTGTGAGGACCTGAGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1941
DB 1965 ATGGCATCCAGTGCAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2024
QY 1942 CCAGCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2001
DB 2025 ACACCGGAGAACATCGTGTGCAAGAACACAGATCTCTGCTCAACTGCTGCTGCTGCTGCTGCT 2084
QY 2002 CCGGATACAGAGGAGAACAGCCGCTGCGGGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2061
DB 2085 CACGATTCATGCTGCTCAGAGCCCTTCCGCGAGGTGTGCAACATGCCATGGCCAAACAA 2144
QY 2062 AAGTGTGTGAGGCGGCAAAACCCATGCAAGGACAGACACAACTGCCAACAGCAGCGG 2121
DB 2145 AGTGTGTGCAAAACCGGAAACCCCTGCAAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2204
QY 2122 AGTGTATCTACTGGGCACTTCAAGGACCCCATGTACAGTGGAGTGGAGTGGAGTGGAGTGG 2181
DB 2205 AGTGTCACTACTGGGCTCACTACAGTACCCCATGTACCGCTGTGAGTGGAGTGGAGTGGAG 2264
QY 2182 ACAGGCGGACAGGCGCTCATCTGCGGGGAGGACTCGGACCTCGGAGCTGGAGCTGGCCCAAC 2241
DB 2265 ATGACAGCAATGACATCATCTGCGGAGAGGACACAGACCTGAGCGGCTGGCGCTATGAA 2324

```

2242 ATCTGGTCTGGCCACCAACCGCCTACCTACCTGATCATCAAGGATACCTGCCCCATCTGC 2301
2325 ACCTGGTGTGTGGCCCAACCGCACTACCTGCAAAAAGGCAACTGCCCCCACTTC 2384
2302 CAATTTGGCCAGGAAGACTTTGCAAGAGCGGATTCGGGATGCGCTGTGATGAGCG 2361
2385 CCACTGGGGCAGGAAGACTATGCAAGAGCGGATTCGGGATGCGCTGTGATGAGCG 2444
2362 ATGACAATGAGGGTGTGACCGATGAGAGGACAACTGCCAGCTCCTCTCAATCCCGCC 2421
2445 ATGACAAGCAAGATCCCAGATGACGGGACAACTGTCTCATTTCAATTAACCCAGCCC 2504
2422 AGGCTGACTATGACAAGGATGAGTTGGGGACCGCTGTGCAACTGCCCTTACGTGCACA 2481
2505 AGTATGACTATGACAGAGATGATGTGGGAGACCGCTGTGCAACTGCCCTTACAAACCA 2564
2482 ACCCTGCCAGATGCACACAGACACAACTGAGAGGGTGACCGCTGCTCGTGAGCATTCG 2541
2565 ACCCTGACCAAGACAGACACAAAGACGGGAGGGCGATGCTGTGTGTGGACATCG 2624
2542 ATGGGGACGATGTCTTCAATGAACAGAGACAAATGTCCTTACGCTCTACAACTGACCAGA 2601
2625 ATGGAGATGGAATCTCAATGAACAGAGACAACTGCCAGTACGTTTACAAGCTGACCA 2684
2602 GGGACAGGATGGTGACCGTGTGGGGATCACTGTGACAACTGCCCTGCTGCACAAAC 2661
2685 GGGACAGGATGGTGATGGGGTGTGGATGACGTGTGACAACTGCCCTGGAACACAATC 2744
2662 CTGACCAAGACCGAGTGACAACTGCTTGTGGGACCAAGTGTGACAAACAGGAGACA 2721
2745 CAGACAGCTGACTCTGACTCAGACCTCATAGGGGACACTTGTGACAACTCAGGACA 2804
2722 TAGATGACGCGGCACAGACACACAGGACAACTGCGCCCTATCTCCAGCGCAAC 2781
2805 TCGATGAGGATGGCCATCAGAACAACTCTGGACAACTGTCCCTATGTCCTTAATGCCAAC 2864
2782 AGGCTGACCATGACAGACAGCGCCAGGCGAGCGCTGTGACCCCTGATGATGACAAAGATG 2841
2865 AGGCGACCATGATAAGATGCAAGAGAGATGCTGTGACCATGACGATGACAATGACG 2924
2842 CGCTCCCGATGACAGGACAACTCGCGCTGTGTGTCAACCCAGACAGGAGACTTGG 2901
2925 GCATCCCTGATGACAGACAACTCGAGCGCTGTGCGCAATCTGACCAAGAGACTCTG 2984
2902 ACGTGTGAGCGGGGTGATATTTGTAAGATGATTTTGACATGATGACAACTCCAGATA 2961
2985 ATGTGATGGCGAGTGACGCTGCAAGACGACTTTGACCATGACATGTGCCAGATA 3044
2962 TTGATGATGTGTCTGAAACAAATGCCATCAGTGAGACAGACTTCAGAACTTCAGA 3021
3045 TTGATGACATCTGTCTGAGAAATTTGACATCAGTGAAACCGGATTTCCGACAAATTCAGA 3104
3022 TGGTCCCTTGGATCCCAAGGAGCACCAAAATTTGATCCCACTGGTCAATTCGCCATC 3081
3105 TGATCTCTAGATCCCAAGGAACTCCCAAAATGACCTTAAGCTTGGTGTCCGCCATC 3164
3082 AAGCAAGAGACTGTTTCAGACAGCAAACTCGAGCCCGGATCGCTGTAGGTTTGAAG 3141
3165 AGGCAAGAACTTGTCCAGACTGTAAACTGTGACCCCTGAGCTGTGTGTAGGTTATGATG 3224
3142 AGTTTGGTCTGTGACTTCACTGACATTTCTACGTAAACCTGACCGGAGCAGCACT 3201
3225 AGTTTATGCTGTGACTTTCAGCGGTACCTTCTCATCAACCCGAAAGATGATGACT 3284
3202 ATGCCGGTCTGCTTTGGTTACAGTCAAGCAGCGCTTCTATGTGGTGATGTGGAAGC 3261
3285 ACGTGGCTTTGTTTTCGGCTACCACTGACGACGCGCTTCTACGTGTGTGATGTGGAAC 3344
3262 AGGTGACGAGACTTACTGGGAGGACCCAGCGGGCTATGCTACTCCGGCTGT 3321
3345 AAGTCAACGACTTCTACTGGGACCAACCCCAAGGGCTCAGGATATCTCAGGCTGT 3404
3322 CCTCAAGTGTGTGAATCTCCACACCGGAGCGGCGAGCACTTGAGGAACCGCTGTGGC 3381

Db 3405 CTGTAAAGTGTGTGAATCTCCACCGCCCTGGCGAGCACCTGCGGAATGCACTGTGGC 3464
Qy 3382 ACACGGGAAACACGCGGGGAGGTGCGAACCTTATGGCACGACCCAGAGAACTTGGCT 3441
Db 3465 ACACAGGAAACACCGCTGCGCAGGTGCGCACCTGTGGCATGACCTCGCCACATCGCT 3524
Qy 3442 GGAAGGACTACACGGCCTATAGTGGTGGCACCTGACTCACAGGCCCAAGACTTGGCTTACATCA 3501
Db 3525 GGAAGATTTCACTGCATACAGATGGCGTCTCAGCCAGGCCAAAGACCGGTTATATCA 3584
Qy 3502 GAGTCTTAGTGCATGAAGAAACAGGTTCATGGCAGACTCAGGACCTATCTATGACCAAA 3561
Db 3585 GAGTGGTGTATGATGAAGAAAGAAATCATGGCTGACTCGGAGCCATCTATGACAAA 3644
Qy 3562 CCTACGCTGGGGGGCGCTGCTTATTTGTCTTCTCTCAAGAAATGGTCTATTTCTCAG 3621
Db 3645 CCTACGCGGGGTAGACTAGCGCTTGTCTTCTCTCAGGAATGGTGTCTTCTCAG 3704
Qy 3622 ACCTCAAGTACGAATGCAGATATTTAAACAAGATTTGC 3661
Db 3705 ACATGAATACGAGTGTGAGATTCCTATCATCAGCTGC 3744

RESULT 5
AK089315/c

DEFINITION AK089315 3530 bp mRNA linear HTC 20-SEP-2003
LOCUS Mus musculus B6-derived CD11 +ve dendritic cells cDNA, RIKEN
full-length enriched library, clone:F730004F21
product:thrombospondin 1, full insert sequence.

ACCESSION

AK089315.1 GI:26105200

VERSION

HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kishimoto, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Chata, E., Watanabe, K.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 403, 685-690 (2001)

5

The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

TITLE

TITLE

TITLE

TITLE

of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 3530)
 JOURNAL
 REFERENCE
 AUTHORS
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M., and Hayashizaki, Y.
 Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
 Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
 Trust/MRC Building Addenbrookes Hospital Cambridge) whose
 assistance we gratefully acknowledge.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 Location/Qualifiers
 1. 3530
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM_DB.F730004F21"
 /db_xref="MGI:2428221"
 /db_xref="taxon:10090"
 /clone="F730004F21"
 /cell_type="36-derived CD11 +ve dendritic cells"
 /cclone_lib="RIKEN full-length enriched mouse cDNA library"
 misc_feature
 1. 3530
 /notes="thrombospodin 1 (MGI:MGI:98737, GB|NM_011580,
 evidence: BLASTN, 98%, match=2839)"
 ORIGIN
 Query Match 24.9%; Score 1416.6; DB 11; Length 3530;
 Best Local Similarity 69.9%; Pred. No. 5.6e-202;
 Matches 1975; Conservative 0; Mismatches 829; Indels 21; Gaps 4;
 QY 840 ATCAGTGAGACACAGAGACGCTGGCTGGTCCGATGTCACACCGAGTACCTGGGC 899
 DB 3487 ACCCTTTGACAAACACGTTGGTGAAGCTTCCAGCCCTGCTATCCGACCAACTACATCGGC 3428
 QY 900 CCCAGCTCAGAGAGAGGAGCCGAGGTGTGCGAACCGCTCGTGGAGAGCTGGGAACATG 959
 DB 3427 CACAAACAAAGACCTCCAGAGTATCTGTGGCTCTCTCTGTGATGAATATCCAGCATG 3368
 QY 960 GTCCAGAGCTCTCGGGGCTCCAGTCTCTGTGAACAGCTCAGGAGAACCTCAAGAGA 1019
 DB 3367 GTCTTGAATGAAGGGCCTCGGACCATCGTACCACTCTGTGAGGACAGCATCCGAAA 3308
 QY 1020 GTGTCAATGATACCAAGTTTCTCTGGAGGCTCATTTGGTGGCCCTCCCTAAAGACAAGAAC 1079
 DB 3307 GTACGCGAGAGAACAGAGAGCTGTCTAGTGAGCTGAAGCGGCTCCCTCC----- 3257
 QY 1080 ATGTACGCTTGTGGCAGGATGGCGGTTCTTTCGCGAAATGAACAGTGGTGTGGAC 1139
 3256 -----TGCTTTTCAATGGAGTCCAGTACAAGAACACAGAGAGTGGACTGTAGAC 3206
 1140 AGCTGCACACAGCTGTACCTGCACGAATTTTAAACCACTTTGCCACCAATCCTCTGCCG 1199
 3205 AGTTGCACAGAGTGTCACTGCCAGAACTCGGTACCATCTGCAAAAGGTGCTCTGTCTCT 3146
 1200 CTTGCAACCTGCGCCAGTCCATCTTTTGTGGAAGCGAATCTGCTCCCTCTCTCTGCCCTCC 1259
 3145 ATCATGCCCTGTCTCCAAAGCCACAGTTTCTTGTGATGTTGTAATCTGCTGCCACCGTGTGG 3086
 1260 TGGTGGAGCGTGTAGAGAGGCTGGTCTCCCTGGCAGAGTGGACCCAGTCTCTCTGGAGC 1319
 3085 AGCGACTCTGTGACATGGCTGGTCTCCCTGGTGTCTGAGTGGAGCTCTCTCTCTGCCACA 3026
 1320 TGTGGCTCTGGGACCCAGCAGAGAGGCGGTCTCTGTGACGCTCACCAGACAACCTGCTTGTG 1379
 3025 TGTGGCAATGGAATTCAGCAACGTTGGTCTCTGTGACAGCTCAACAACAGATGGAG 2966
 1380 GGGCCCTCCATCAGACACGGGCTTGAGTCTGAGCAAGTGTGACACCGCATCGGAG 1439
 2965 GGCTCTTCGGTACAGACAAGGACCTGCCAATTGAGGAGTGTGACAAAGATTTAAACAG 2906
 1440 GACGGCGGTGGAGCCACTGGTCACTTGGTCTTTCATGCTCTGTGACCTGTGGAGTGGC 1499
 2905 GATGGTGGTGGAGTCACTGGTCTCCATGGTCTGCTCTGTGACCTGTGGTGGAGT 2846
 1500 AATATCACAGCATCGTCTCTGCAACTCCCACTGCCCCAGATGGGGGGGAGAAATTCG 1559
 2845 GTGATCACAGGATTCGGGCTCTGCACTCCCCAGCCCCAGATGAACGGGAAGCCCTGT 2786
 1560 AAGGGAGTGGCGGGAGACCAAGCCCTGCCAGGGCGCCCATGCCAATCGATG-CG 1618
 2785 GAAGGTGAAGCCCGGGAGACCAAGCCCTGCAAGAAAGACGCTGCCCAATTAATGGAGG 2726
 1619 CTGGAGCCCTGTCTCCCTGGTCTGGGCTGGCTGCACTGTGACCTGTGCGGTGGATCCGGA 1678
 2725 CTGGGTCTCTGTGCTACCATGGGACATCTGCTCTGTGCTACCTGTGGAGGAGTGCAGAG 2666
 1679 GCGCACCGGCTCTGCAACAGCCCTGAGCTCAGTACGGAGGAGAGGCGCTGGTGGGGGA 1738
 2665 ACCAGCCGACTCTGTAAACAACCCACACCCAGTTTGGAGGCAAGAGCTGTGTTGGCGA 2606
 1739 TGTG-CAGGAGCGTCAATGTGCAACAGAGGAGTGGCCCGTGTGATGGCTGTTATCCA 1797
 2605 TGTGACAGAAAAATCAAGTTTGGCAAGCAGGAGTGGCCCAATTGATGGATGCTGTCCA 2546
 1798 ACCCTT-GCTTCCGGGAGCCAGTGCAGCAGCTTCCCGGATGGTCTCTGTGCTCATCGGC 1856
 2545 ATCCCTGGCTTGTGTTGGTCCCAAGTGTACTAGTACCTGATGGTGGAAATGTGT 2486
 1857 TCCTGCCCTGTGGGCTTTCTTGGGCAATGGCAACCACTGTGAGGACCTGGACAGTGTGC 1916
 2485 GCGTGTCTCTCTGGCTACAGTGAATGGCATCCAGTGCNAAGAGCTGATGAGTGCAA 2426
 1917 CTGGTCCCCGACATCTGCTTCTCCACCAAGAGTGGCTCTGCTGTGTCAACACTCAGCCT 1976
 2425 GAAGTGGCTGATGCTTGTCTTCAATCAACAGGAGAACATCGTGCAGAACACAGATCCT 2366
 1977 GGCTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2036
 2365 GGCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2306
 2037 CTGGAAGCAGCAAGACGGAAAGAGTGTGTGAGCCCGGCAAAACCCATGCAAGGACAAG 2096
 2305 GTCGACATGCCATGGCCCAACAAACAGTGTGTGCAACCGCGAAACCCCTGCGACGAGGG 2246
 2097 ACACAACTGCAACAGCAGCGGAGTGTGATCTATCTGGGCGCACTTACGACACCCCATG 2156
 2245 ACAGTGTGCAACAGAGAGCTTAAAGTGAACACTCTGGGCTCACTACAGTACCCCATG 2186
 2157 TACAAGTGGCAGTGCAGAGAGGCTTACGGCGGACGCGGCTCATCTGCGGGGAGGAGTCTG 2216
 2185 TACCGCTGTGAGTGAAGCCCGGCTATGACAGGCAATGSCATCATCTCGGAGAGGACACA 2126

2217 GACCTGGACGGCTGCCCAACCTCAATCTGRTCTGCGCCACCAACGCCACCTACCACTGC 2276
 2125 GACCTGGACGGCTGCCCTAAATGAAACCTTGGTGTGTGGCCAAACCAACCTACCACTGC 2066
 2277 ATCAGGATTAAGTCCCTGCTGCAATCTGGCAGGAGACTTTGACAGGAGCGG 2336
 2065 AAAAGGACAACTGCCCAACCTTCCCACTCGGGCAGAGACTATGACAGGAGCGG 2006
 2337 ATTGGCGATGCTGTGATGATGACGATGCAATGACGGTGTGACCGATGAGAGGCAAC 2396
 2005 ATTGGCGATGCTGTGATGATGACGATGCAACGACGATCCCGATGACAGGAGCAAC 1946
 2397 TGGCAGCTCTTCAATCCCGCCGAGGCTGACTATGACAGGATGAGTTGGGACCGC 2456
 1945 TGTCCATTCATTAACCCAGCCAGTATGACTATGACAGAGATGATGTGGAGACCGC 1886
 2457 TGTGACAACTCCCTTACGTGACCAACCTGCCAGATCGACACAGCAACAAATGGAGAG 2516
 1885 TGTGACAACTGCCCTTCAACCCACCAACCTGACCAAGCAGACACAGCAACAAACGGGAG 1826
 2517 GGTGACGCTCTCGTGGACATTTGATGGGACGATGCTTCAATGAAAGGACAAATGT 2576
 1825 GGCATGCTGTGTGGACATCGATGAGATGGAATCCTCAATGAACGAGACAACTGC 1766
 2577 CCCTACGCTCTACACACTGACAGAGGACACGGATGGTCAACGGTGTGGGGATCACTGT 2636
 1765 CAGTACGTTTACACGCTGGACAGAGGACACGGACATGGATGGGTGGAGATCAGTGT 1706
 2637 GACAACTGCCCTGCTGACACACCTGACACCGACGAGCGGATGACAAATGACCTTGTGGG 2696
 1705 GACAACTGCCCTGCTGACACCAATCGACACCGAGTGGATCTGACTCGACACCTCATAGG 1646
 2697 GACCACTGTGACAAACAGGAGGACATAGATGACGAGCGGACACAGCAACCAAGGACAA 2756
 1645 GACACTTGTGACAAACATCAGGACATCGATGAGATGGCCATCAGAACAACTCTGGACAC 1586
 2757 TGCCCTACATCTCAACGCCAACACGGCTGACCATGACAGAGACGGCCAGGCGAGCC 2816
 1585 TGTCCTATGTGCTTAATGCCAACAGGCGGACCATGATTAAGATGGCAAGGAGATGCC 1526
 2817 TGTGACCTGATGACAAACGATGGCTGCCGATGACAGGAGCAACTGCCGGCTGTG 2876
 1525 TGTGACCATGACGATGACATGACGGCATCTCTGATGACAGAGCAACTGCGAGCTGGT 1466
 2877 TTCAACCCAGACAGGAGACTTGGACGGTGTGATGACGGGGTATATTTGTAAGATGAT 2936
 1465 CCAATCTGACAGAGGACTCTGATGATGATGGCCGAGGTGACGCTGCAAGAGCGAC 1406
 2937 TTTGACATGACAAACATCCAGATATTGATGATGTGTCTGCTGAAACAAATGCCATCAGT 2996
 1405 TTTGACCATGACATGTGCCAGATATTGATGACATCTGCTGAGATTTTGACATCAGT 1346
 2997 GAGACAGACTTCCAGAACTTCCAGATGTCCTGATGATGATGATGATGATGATGATGAT 3056
 1345 GAAACCGATTTCCGACAAATTTCCAGATGATTTCTAGATCCCAAGGAACTCCCAAAAT 1286
 3057 GATCCCACTGGTCTATTCGCCATCAAGCAAGGAGCTGTTGTCAGACAGCAACTCGGAC 3116
 1285 GACCTTAAGTGGTGTGCGCCATCAGGCAAGAACTTGTCCAGACTGTAACTGTGAC 1226
 3117 CCCGGCATCGTGTAGTTTTCAGGATTTGGTGTGTGAGCTTCACTGAGTCACTTCTAC 3176
 1225 CCTGGACTTGTGTAGTTTATGATGATTTAATGCTGTGAGCTTCACTGAGCTTCTTCTTC 1166
 3177 GTAACACTGACCGGAGGACGACACTATCGCGCTTCTGTTGGTTACAGTCAAGCAGC 3236
 1165 ATCAACCCGAGAGATGATGACTACGTCGCTGCTTGTGCTGCTGCTGCTGCTGCTGCT 1106
 3237 CGCTTCTATGTGGTGTGATGAGAGGAGTGTGACGACACTTCTGAGGAGCCAGCCACG 3296
 1105 CGCTTCTACGTGTGATGTGGAGACAAAGTCAACCCAGTCTCTGAGACACCAACCCACCA 1046

QY 3297 CGGCGCTATGGCTACTCCGGCGTGTCCCTCAAGTGTGTGAATCTCCACCGGGACGGC 3356
 Db 1045 AGGCTCAGGGATACTCAGCGCTGTCTGTAAAGTTGTAACTCCACCGGGCCCTGGC 986
 QY 3357 GAGCACCTGAGAAACGCGCTGTGGCACACGGGAAACACCGCGGAGGTGCGAACCTTA 3416
 Db 985 GAGCACCTGCGGATGCACTGTGGCACACAGGAAACACCGCTGCGGAGGTGCGAACCTG 926
 QY 3417 TGGCAGCACCCAGGAAACATTTGGCTGGAGGACTACACGGCCTATAGTGGCACCTGACT 3476
 Db 925 TGGCATGACCTCGCCACATCGCTGGAAAGATTTCACTGCATACAGATGCGCTCTCAGC 866
 QY 3477 CACAGGCCAAGACCTGCTACATCAGAGTCTTAGTGCATGAGGAAACAGGTTCATGSCA 3536
 Db 865 CACAGGCCAAGACCGGTTATATCAGAGTGTGATGATGAGGAAAGAAATCATGCTC 806
 QY 3537 GACTCAGGACCTATCTATGACCAACCTACGCTGGCGGGCGCTGCTCTATTGTCTTC 3596
 Db 805 GACTCGGACCATCTATGACAAACCTACGCGGGGTAGACTAGGCTGTTCGCTTC 746
 QY 3597 TCTCAGAAATGGTCTATTCTCAGACCTCAAGTACGATGCAAGATATTTAAACAAGA 3656
 Db 745 TCTCAGAAATGGTCTTCTCTCAGACATGAATACGAGTGTGAGATTCCTAATCATCA 686
 QY 3657 TTTGC 3661
 Db 685 GCTGC 681

RESULT 6
 AY405252
 LOCUS
 DEFINITION Homo sapiens THBS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCSSION AY405252
 VERSION AY405252.1 GI:39761226
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 3412)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 3412)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment
 FEATURES
 Location/Qualifiers
 1..3412
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>3412
 /gene="THBS1"
 /locus_tag="HCM2162"

gene
 source
 ORIGIN

Query Match 23.9%; Score 1360.8; DB 29; Length 3412;
 Best Local Similarity 61.2%; Pred. No. 1.2e-193;
 Matches 2142; Conservative 0; Mismatches 1243; Indels 114; Gaps 3;


```
1074 AACTCAGGGCAGGAAGACTATGACAAGATGGAATTGGTGATGCGCTGTGATGATGACGAT 2133
1075 GACAAATGACGGTGTGACCGATGAGAAGACAACTGCCAGCTCTCTTCAATCCCGCCAG 2423
1076 GACAATGATAAAATTCAGATGACAGGGACAACTGTCCATTCCATTACAACCCAGCTCAG 2193
1077 GCTGACTATGACAGAGTGAAGTTGGGACCGCTGTGACAACTGCCCTTACGTGCAAC 2483
1078 TATGACTATGACAGAGATGATGTGGGACCGCTGTGACAACTGTCCCTTACAAACCAAC 2253
1079 CTTGCCAGATTCGACACAGACAACTGAGAGGGTGAAGCTGTGTCCTGCGTGAACATTGAT 2543
1080 CCAGATCAGGACAGACACAGACAACTGAGAGGGAGAGCGCTGTGCTGCAGACATTGAT 2313
1081 GGGGACGATGCTTCAATGACAGAGACAACTGCTCCCTAGCTGTACAACTGACAGAGG 2603
1082 GGAACCGGTATCTCAATGAAACGGGACAACTGCCAGTACGCTTACAATGTGACCAAGAG 2373
1083 GACACGGATGCTGACGGTGTGGGGATCACTGTGACAACTGCCCTGTGTCACAACTT 2663
1084 GACACTGATATGATGGTGGTGGAGATCAGTGTGACAACTGGCCCTTGGAAACACATCCG 2433
1085 GACAGACCGAGCTGGACAACTGCTTGTGGGACCGAGTGTGACAACTGAGGAGCAATA 2723
1086 GATCAGCTGACTCTGACTCAGACCGCAATTGGAGATACCTGTGACAACTCAGGATATT 2493
1087 GATGACACGGCCACCAAGACAACTGAGGACAACTGCCCTTACATCTCCAACTGCAACCAAG 2783
1088 GATGAAGATGGCCACCAAGACAACTGAGGACAACTGCTCCCTATGTGCCAATGCCAACCG 2553
1089 GCTGACCATGACAGAGACGGCGAGCGGCGAGCTGTGACCTGTGATGATGACAACTGATGC 2843
1090 GCTGACCATGACAAAGATGGCAAGGAGATGCTGTGACCGAGATGATGACAACTGATGC 2613
1091 GTCCCGATGACAGGACAACTGCGGCTGTGTTCAACCCAGACAGGAGACTTGGAC 2903
1092 ATTCCTGATGACAGGACAACTGCACTGTGCGCAATCCGACAGAGGACTCTGAC 2673
1093 GGTGATGACAGGGGTGATATTGTAAGATGATTTGACAACTGACAACTCCAGATATT 2963
1094 GGGATGCTGAGGTGATGCTGCAAGATGATTTGACCATGACAGTGTGCCAGACATC 2733
1095 GATGATGTGTCTGAAACAACTGCCATCAGTGAGACAGATTCAGGAACTCCAGATG 3023
1096 GATGACATCTGCTGAGATGTTGACATCAGTGAGACCGATTCGCGGATTCAGATG 2793
1097 GTCCCTTGGATCCCAAGGACCACTCAAAATGATTCCTCAACTGGGTCTATTCGCCATCAA 3083
1098 ATTCCTCTGGACCCCAAGGACATCCCAAAATGACCTTAACCTGCTGAGGTTTACGCCATCAG 2853
1099 GGCAGGAGCTGTTTACAGACAGCACTCGGACCCGGCATCGCTGATGTTTACGAG 3143
1100 GGTAAAGAACTCGTCCAGCTGTCACTGTGATCTGCTGAGTCTGCTGATGATGAG 2913
1101 TTTGGGCTGTGGACTTCAGTGGCACTTCACTGTAACACTGACCGGACGACGACTAT 3203
1102 TTTAATGCTGTGGACTTCAGTGGCACTTCTTCACTCAACACCGAAGGACGATGACTAT 2973
1103 CGCGCTTGTCTTGTGTTACAGTCAAGCAGCGCTTCTATGTTGTTGATGTTGGAAGCAG 3263
1104 GCTGGATTTGTCTTGTGGTACCAAGTCCAGTCCAGCGCTTTTATGTTGATGTTGGAAGCAA 3033
1105 GTGACCGAGACCTACTGCGGAGGACCAAGCCAGCGGCTGTGCTTACTCCGCGGTGTC 3323
1106 GTACCCAGCTCTACTGCGGACCAACCAACCCACGAGGGCTCAGGGATCTCTCGGCGCTTCT 3093
1107 CTCAGGTGTGTAATCCACCAAGGAGCGGAGACCTGAGGAACCGCTGTGGCAC 3383
1108 GTGAAAGTGTAAATCTCCACCAAGGCGCTGCGAGCACTCGGAAACCGCTGTGGCAC 3153
1109 ACGGGGACACCGCGGCGAGTGTGAACTTATGACAGACCCCAAGGAACTTGTGCTG 3443
```

```
Db 3154 ACAGGAACACCCCTGGCCAGGTGGCCACCTGTGTGCATGACCCCTCGTCACATAGGCTGG 3213
Qy 3444 AAGGACTACACGGCTATAGGTGGCACCTGACTCAGAGCCCAAGACTGCTCATACAGAG 3503
Db 3214 AAGATTTCACCGCTACAGATGGCGTCTCAGCCACAGCCCAAGACGGGTTTCATTAGA 3273
Qy 3504 GTCTTAGTGCATGAAGGAACAGGTCTATGGCAGACTCAGGACCTATCTATGACCAAC 3563
Db 3274 GTGGTGTATGATGAAGGAAGAAATCATGGTGTGACTCAGGACCCCATCTATGATAAAAC 3333
Qy 3564 TACCTCGCGCGCGCTGGGTCTATTGTCTTCTCTCAAGAAATGGTCTATTTCTCAGAC 3623
Db 3334 TATGCTGGTGTAGACTAGGTGTTGTCTTCTCTCAAGAAATGGTGTCTTCTCTGAC 3393
Qy 3624 CTCAGTACGAATGCAGAG 3642
Db 3394 CTGAATACGAATGTAGAG 3412

RESULT 7
AY405254 3240 bp DNA linear GSS 16-DEC-2003
LOCUS
DEFINITION Mus musculus THBS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
Genomic survey sequence.
ACCESSION AY405254
VERSION AY405254.1 GI:39761228
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3240)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J.,
Adams,M.D. and Cargill,M.
ADAMS M.D. and Cargill M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3240)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J.,
Adams,M.D. and Cargill,M.
ADAMS M.D. and Cargill M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..3240
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>3240
/gene="THBS1"
/locus_tag="HCM2162"

Query Match 21.4%; Score 1219; DB 29; Length 3240;
Best Local Similarity 61.4%; Pred. No. 1.9e-172;
Matches 2071; Conservative 0; Mismatches 1086; Indels 216; Gaps 3;

Qy 270 GGGCCGACCCCGGGGTCCGGCTTACCGCTTTCGTCGCTTTTCTACTATCATCCCAACCGGTG 329
Db 84 GGCCAAGATCTATCCAGCCCGCTTCCGATGAGATGCCAACCCTGATCCCGCTGTG 143
Qy 330 AACCGAGATGACCTCAGCAAGATCACCAGATCATCGGAGAGAGGCGCTTCTTCTC 389
Db 144 CCGGATGACAAGTTCGAAGACCTACTGGACGCTGTGTGGCGCCGACAAAGCTTCATCTTC 203
Qy 390 ACGGCCAGCTCAGCAGGACGGCAAGTCCAGGGGACGCTGTGTGCTCTGAGGGCCCC 449
```



```
2610 GATGGTACGGTGTGGGGATCACTGTGACACTGCTGCCCTCCCTGGTGACACACCTGACCAG 2669
2620 GACATGATGGGTGGAGATCAGTGTGACAACTGCCCCCTGGAAACAAATCCAGACCAG 2667
2670 ACCGACGTGGCAATAGACTTGTGGGACCACTGTGACAAACAGGACATAGATGAC 2729
2680 CTGACCTCTGACTCAGACCTCATAGGGACACTTGTGACAAATCAGGACATCGATGAG 2327
2730 GACGGCCACCAGAACCAACAGGACAACTGCCCTCATCTCCAAACCCACAGGCTGAC 2789
2820 GATGGCCATCAGACAACTGTGGACAACTGCTCCATGTGCTTAACCCACAGGCGGAC 2387
2790 CATGACAGACAGCGCCAGGCGGACGCTGTGACCACTGTGATGACAAACAGTGCCTCCC 2849
2880 CATGATAAGATGCAAGAGATGCTGTGACCACTGACGATGACAAATGACGCGCATCCCT 2447
2850 GATGACAGGAGCAACTGCGGCTGTGTTCACCCAGACAGGACACTTGGACGGTGAT 2909
2448 GATGACAGACAACTGCGGCTGTGCTCCATCTGACCAAGGACTCTGATGTGAT 2507
2910 GGACGGGTGATATTGTGAAGATGATTTGACATGACAACTCCAGATATTGATGAT 2969
2508 GCGCGAGTGACGCGCTGCAAGAGACACTTTGACCAATGACAAATGTGCCAGATTTGATGAC 2567
2970 GTGTGTCTCTGAAAAAATGCACTCAGTGAGACAGACTTCAGGAACCTTCAGATGTTGGG 3029
2568 ATCTGTCTGAGATTTTGACATCAGTGAACCGATTTCCGACGATTCAGATGATTCCT 2627
3030 TTGATCCCAAGGACCAACCAATGATCCCACTGGGTATTCGCCATCAGGCAAG 3089
2628 CTGATCCCAAGGACCACTCCCAATGACCTTAACCTGCTGTTGCTCGCCATCAGGCAAA 2687
3090 GAGCTGTTTCCAGACAGCAACTCGGACCGGCTATCGCTGTAGTTTGGACAGTTTGGG 3149
2688 GAATCTGCTGAGACTGTAACTGTGACCTGCTGCTGATGATGATGATTTAAT 2747
3150 TCTGTGACTTCACTGGACACTTCTACGTACACTGACCGGACGACACTATGCCGC 3209
2748 CTTGTGACTTCACTGGGACTTCTTCTATCAACCGGAGAGATGATGACTGCTGCG 2807
3210 TCTGTCTTGTGTTACCACTGACGCGCTTCTATGTTGATGTTGGAAGCAGGTGACG 3269
2808 TTTGTATTGCTGCTACCTGCTGACGAGCGCTTCTACGTTGATGTTGGAACAAAGTCAC 2867
3270 CAGACCTTCTGGGAGGACCGCCAGCGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3329
2868 CAGTCTCTGGGACCAACCCCAAGGCTCAGGATATCTGAGGCTGCTGCTGCTGCTGCTGCT 2927
3330 GTGTGAACTCCACACCGGACCGGCGGACCTGAGGAACCGCTGTCGACACCGGG 3389
2928 GTTGTGAATCCACACCGGCTGCGGACGACCTGCGGAATGCACTGTGGACACAGGA 2987
3390 AACACCGCGGCGGAGGTCGCAACTTATGGACGACCCCGGACCACTTGGCTGGAGGAC 3449
2988 AACACCTCTGCGGAGTGGCACCTGTTGGGATGACCTCGCCACATCGCTGGAAGAT 3047
3450 TACACGCGCTATAGGTGCACTGACTCAGCGGCGGACGCTGCTGCTGCTGCTGCTGCTGCT 3509
3048 TTTCACTGCTACAGATGCTGCTCAGCCACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 3107
3510 GTGCATGAAGGAAACAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3569
3108 ATGTATGAAGGAAAGAAATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3167
3570 GCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3629
3168 GCGCGTGAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3642
3630 TACGAATGACAG 3642
3228 TACGAGTGTGAG 3240
```

```
RESULT 8
BQ933301
LOCUS
DEFINITION
AGENCOURT_8744271_Lupski_sciatic_nerve_mRNA linear EST 21-AUG-2002
IMAGE:6205936 5', mRNA sequence.
ACCESSION
BQ933301
VERSION
BQ933301.1 GI:22348684
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (Bases 1 to 939)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM13629 row: j column: 17
High quality sequence stop: 686.
FEATURES
Location/Qualifiers
1..939
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6205936"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski sciatic nerve"
/notes="Vector: pCMV-Sport6 (Life Technologies); Site: 1.
Noti; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACCGGTCG-3' and
5'-GACATGTTCTAGATCGGACGGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
```

ORIGIN

```
Query Match 15.0%; Score 854; DB 13; Length 939;
Best Local Similarity 97.6%; Pred. No. 9.1e-118;
Matches 919; Conservative 0; Mismatches 17; Indels 6; Gaps 5;

QY 4706 CTAGTAACACATATGTAATATTTCTAGAACAAACATCTTAATAGTATATATCTG 4765
Db 1 CTAGTAACACATATGTAATATTTCTAGAACAAACATCTTAATAGTATATATCTG 60
QY 4766 TGAATAATATGAGGCTTGATAATATTAGGTTGTCAAGCATCTAGAACCTGTAA 4825
Db 61 TGAATAATATGAGGCTTGATAATATTAGGTTGTCAAGCATCTAGAACCTGTAA 120
QY 4826 AGATACATAGATATATGAGGAGTTTATGATGACCTTAATATATATATCTG 4885
Db 121 AGATACATAGATATATGAGGAGTTTATGATGACCTTAAATATATATCTG 179
QY 4886 CGATTTAGTTCAATATTTGTTACTGTTATCTATCTGTTATATGGAATCTTTAAT 4945
Db 180 CGATTTAGTTCAATATTTGTTACTGTTATCTATCTGTTATATGGAATCTTTAAT 239
QY 4946 CAACGCTGAAA-GAATCAGCATTTAGTCTGCCAGGCACACCCCAATATCATCATGT 5004
```



```

661 ATGCTTAGTACACATATGATGATATATTTCTAGAAACAAACATCTAATAAGTATATAT 720
4762 -CCTGTGAAATATAGAGCTTGTGATAATATTAGGTTGTCTACGATGAAGCATGCTAGAAGCT 4820
721 CCCTGTGAAATATAGAGCTTGTGATAATATTAGGTTGTCTACGATGAAGCATGCTAGAAGCT 780
4821 GTAACAGATACATAGA-GAATAATGAGGAGTTTATGAT-GGAACCTTAAATATATAATG 4878
781 GTAACCGATACATAGAGGAATATGAGGAGTTTATGATGCGGACCTTAAATATATAATG 840
4879 TT 4880
841 GT 842

RESULT 10
BX390187 896 bp mRNA linear EST 08-MAY-2003
LOCUS BX390187 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DC020YK24 5-PRIME, mRNA sequence.
ACCESSION BX390187
VERSION BX390187.1 GI:30461404
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 896)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: sequef@genoscope.cns.fr Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen corporation 1800
Faraday Avenue Genoscope sequence ID : CS0BAG0072F12_CS00670_1.
Location/Qualifiers
1..896
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC020YK24"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-cligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source
1..896
Query Match 13.9%; Score 790.2; DB 13; Length 896;
Best Local Similarity 97.0%; Pred No. 3.1e-108;
Matches 835; Conservative 0; Mismatches 22; Indels 4; Gaps 3;

QY 138 TGGAGCTGGTCTGCTGGCTCTGTGGG--TGTGGCCAGACGACGAGCTGGTCAACGAG 195
DB 1 TGTGTGCTGGTCTGCTGGCTCTGTGGGTTGTGGGCTGACGACGACGAGCTGGTCAACGAG 60

QY 196 ACAAGACAGCAGCTTCGACCTTTTCAGTATCAGACATCAACCGCAAGACCATTTGGCG 255
DB 61 ACAANACAGCAGCTTCGACCTTTTCAGTATCAGACATCAACCGCAAGACCATTTGGCG 120

QY 256 CAAGCAGTTCCGGGGCCGACCCGGGCTGCGGCTTACCGTTTGGTGGCTTGGACT 315
DB 121 CAAGCAGTTCCGGGGCCGACCCGGGCTGCGGCTTACCGTTTGGTGGCTTGGACT 180

QY 316 ACATCCACCGGTGAACGACAGATGACCTCAGCAGAGATCACCAGATCATGCGGAGAGG 375
DB 181 ACATCCACCGGTGAACGACAGATGACCTCAGCAGAGATCACCAGATCATGCGGAGAGG 240

```

```

QY 376 AGGCTTCTTCTCAGCGCCAGCTCAGCAGAGCGGCAAGTCCAGGGGACGCTGTGG 435
DB 241 AGGCTTCTTCTCAGCGCCAGCTCAGCAGAGCGGCAAGTCCAGGGGACGCTGTGG 300
QY 436 CTCTGGAGGGCCCGGTTCTTCCAGAGGAGGTTTCAGATCGTCTCCAAAGGCCCGCGG 495
DB 301 CTCTGGAGGGCCCGGTTCTTCCAGAGGAGGTTTCAGATCGTCTCCAAAGGCCCGCGG 360
QY 496 ACAGCTGGATCTCAGCTTACTTGGATTGACGGACCCCGGATGTTGGTCTCCCTGGAGGAGC 555
DB 361 ACAGCTGGATCTCAGCTTACTTGGATTGACGGACCCCGGATGTTGGTCTCCCTGGAGGAGC 420
QY 556 TCGGCTGGCTGACTCGCAGTGGAGAGAGCTACCGTGCAGGTGGCTGGCGAGACCTACA 615
DB 421 TCGGCTGGCTGACTCGCAGTGGAGAGAGCTACCGTGCAGGTGGCTGGCGAGACCTACA 480
QY 616 GCTTCAGCTGGGCTGGCGACCTCATAGACAGCTTCGCTCTTGGACGAGCCCTTCTACGAGC 675
DB 481 GCTTCAGCTGGGCTGGCGACCTCATAGACAGCTTCGCTCTTGGACGAGCCCTTCTACGAGC 540
QY 676 ACCTGCAGCGGAAAGAGCGGATGTAGTGGCCAAAGGCTCTGCCAGAGAGAGTCACT 735
DB 541 ACCTGCAGCGGAAAGAGCGGATGTAGTGGCCAAAGGCTCTGCCAGAGAGAGTCACT 600
QY 736 TCAGGGGTTTGGCTTCAGAACGTCCTACCTAGTTGTTGAAAACCTCTGTGGAAGATATTCTAA 795
DB 601 TCAGGGGTTTGGCTTCAGAACGTCCTACCTAGTTGTTGAAAACCTCTGTGGAAGATATTCTAA 660
QY 796 GCAAGAAGGTTGCCAGCAAGGCCAGGAGCTCAGATCAACGCCATCAGTGAGAACACAG 855
DB 661 GCAAGAAGGTTGCCAGCAAGGCCAGGAGCTCAGATCAACGCCATCAGTGAGAACACAG 720
QY 856 AGACCTGCGCTGGTCCGATGTCCACACGAGTACGTGGGCCCCCAGCTCAGAGAGA 915
DB 721 AGACCTGCGCTGGTCCGATGTCCACACGAGTACGTGGGCCCCCAGCTCAGAGAGA 780
QY 916 GSCCCGAGTGTGCGAACCGCTCGTGCAGAGAGCTGGGAAACATCGTCCAGAGAGTCTCGG 975
DB 781 GSCCCGAGTGTGCGAACCGCTCGTGCAGAGAGCTGGTGCAGAGAGTCTCGGAGCTC-G 838
QY 976 GSCCTCACGCTCTCTGTGAACC 996
DB 839 GSCCTCACGCTCTCTGTGAACC 859

```

RESULT 11

CA414667 778 bp mRNA linear EST 07-NOV-2002
 UI-H-E20-bar-m-21-0-UI.s1 NCI CGAP_Ch1 Homo sapiens cDNA clone
 UI-H-E20-bar-m-21-0-UI 3', mRNA sequence.

ACCESSION

CA414667
 VERSION
 CA414667.1 GI:24777318

KEYWORDS

EST.
 Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 778)

REFERENCE

NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

AUTHORS

Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

TITLE

Tissue Procurement: Dr. Steven Gitelis/ Rush Presbiterian, Dept. of
 Orthopaedics

JOURNAL

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 DNA Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu

The following repetitive elements were found in this cDNA.

sequence: 356-404, >(CAAA)n#Simple_repeat
Seq primer: M13 FORWARD
POLYA=tes.

FEATURES

Location/Qualifiers
1..778
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-E20-bar-m-21-0-UI"
/tissue_type="Chondrosarcoma Grade II"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP-Ch1"
/note="Organ: Left Pelvis; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ch1 is a cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCGT.
TAG TISSUE=grade-2-chondrosarcoma
TAG LIB=UI-H-E20
TAG_SEQ=ACTTAATATG"

ORIGIN

Query Match 13.3%; Score 760; DB 14; Length 778;
Best Local Similarity 100.0%; Pred. No. 1.1e-103;
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4644 TAATATTTTAAAGTACATATAGAGCACTCAATTTACAGTAACCTTTTAAATAT 4703
4703 TAATATTTTAAAGTACATATAGAGCACTCAATTTACAGTAACCTTTTAAATAT 719
4704 GCCTAGTAACACATATGTAGTAAATTTCTAGAAACAAACATCTAAATAGTATATCC 4763
4763 GCCTAGTAACACATATGTAGTAAATTTCTAGAAACAAACATCTAAATAGTATATCC 659
4764 TGTGAAATATGAGGCTTGATATATAGTTGTGCAGTAAGCATGCTAGAGCTGTA 4823
4823 TGTGAAATATGAGGCTTGATATATAGTTGTGCAGTAAGCATGCTAGAGCTGTA 599
4824 ACAGATACATAGAGAAATAGAGAGTTTATGATGGAACCTTAAATATATATGTTGCC 4883
4883 ACAGATACATAGAGAAATAGAGAGTTTATGATGGAACCTTAAATATATATGTTGCC 539
4884 AGCGATTTTGTAGTCAATATTTGTACTGTATCTATCTGCTGTATATGGAATCTTTAA 4943
4943 AGCGATTTTGTAGTCAATATTTGTACTGTATCTATCTGCTGTATATGGAATCTTTAA 479
4944 TTCAACCGCTGAAAGAAATCAGATTTAGTCTTCCAGGCAACCCATATCAGTCAATG 5003
4944 TTCAACCGCTGAAAGAAATCAGATTTAGTCTTCCAGGCAACCCATATCAGTCAATG 419
5004 TGTAAATATGCAAGTTGTTTGTGTTTGTGTTTGTGTTGTTGTTGTTTGTGTTTGTG 5063
5063 TGTAAATATGCAAGTTGTTTGTGTTTGTGTTTGTGTTGTTGTTGTTTGTGTTTGTG 359
5064 CTTTAAGTTCATGATCTTCTGAGAAATAGTCACTCATCCACTCCCATAGGGGT 5123
5123 CTTTAAGTTCATGATCTTCTGAGAAATAGTCACTCATCCACTCCCATAGGGGT 299
5124 TTAGTAAAGAGAGTCTGCTGCTGATGATGATAGGGGGCAAAATCTTTTCCCTTTCT 5183
5183 TTAGTAAAGAGAGTCTGCTGCTGATGATGATAGGGGGCAAAATCTTTTCCCTTTCT 239
5184 GTTAATAGTCATACATTTCTATGCCAAACAGGACCAATCCATTAATAGTCTTAATGT 5243

Db 238 GTTAATAGTCATACATTTCTATGCCAAACAGGAAATCCATTAATAGTCTTAATGT 179
Qy 5244 ACACATTCGATTTTGCATTAATTTTGTGTTTCCCTTGGAGTTGATCGTTGTTG 5303
Db 178 ACACATTCGATTTTGCATTAATTTTGTGTTTCCCTTGGAGTTGATCGTTGTTG 119
Qy 5304 TTGTTTCTGCTGACATTTTACATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 5363
Db 118 TTGTTTCTGCTGACATTTTACATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 59
Qy 5364 CGTTGGGATACCTTCAATTAATGCTAGGACTGTCTCAACAGCG 5403
Db 58 CTTGGGATACCTTCAATTAATGCTAGGACTGTCTCAACAGCG 19

RESULT 12
LOCUS CK000113 763 bp mRNA linear EST 26-NOV-2003
DEFINITION AGENCOURT 16368996 NIH_MGC 220 Homo sapiens cDNA clone IMAGE:30707733 5', mRNA sequence.
ACCESSION CK000113
VERSION CK000113.1 GI:38526147
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 763)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-remail.nih.gov
Tissue procurement: James Martin, University of Iowa
cDNA library preparation: M. Bento Soares, University of Iowa
cDNA library arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>
Plate: NDAM1072 row: m column: 22
High quality sequence stop: 664.

FEATURES

Location/Qualifiers
1..763
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30707733"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC 220"
/notes="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI; Site 2: NotI; Library is oligo-dT primed and directionally cloned. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. Average insert size 0.5-1kb. Adaptors 5'(AATCGCACGAGG)3' and 5'(CTCGTGGCG)3'. 3' linker sequence - GCGGCGGTGAGAGCC T18. Sequencing primers 3' end: T3 promoter primer 5'd (ATTACCTCTCACTAAGGA)3', 5' end: T7 promoter primer 5'd (TAATACGACTCACTAAGG)3'. Library was constructed in the laboratory of M. Bento Soares. Average insert size 3-4kb
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 13.2%; Score 753; DB 14; Length 763;
Best Local Similarity 99.7%; Pred. No. 1.2e-102;


```

Db 541 GTAATAGCCTTTGGTAACGCGGCAAGTTGCTTTTACCTGATTTGATGATACATTTCA 600
Qy 4617 TTAAGTTCCAGTTATAAATATTTTCTTAATATTTTAAAGTGAATGAACTC 4676
Db 601 TTAAGTTCCAGTTATAAATATTTTCTTAATATTTTAAAGTGAATGAACTC 660
Qy 4677 CATTTACAGTAACATTTATTTAAATATGCTAGTAAACACATATGTAATAATTTCTAGA 4736
Db 661 CATTTACAGTAACATTTATTTAAATATGCTAGTAAACACATATGTAATAATTTCTAGA 720
Qy 4737 AACAAATCTTAATAGTATATAATCCTGTGAAATATAGG 4778
Db 721 AACAAATCTTAATAGTATATAATCCTGTGAAATATAGG 762

RESULT 14
BQ940133 904 bp mRNA linear EST 21-AUG-2002
AGENCOURT_8784556 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6376342
5', mRNA sequence.
ACCESSION BQ940133
VERSION BQ940133.1 GI:22355611
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 904)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2558 row: f column: 23
High quality sequence stop: 551.
Location/Qualifiers
1. .904
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6376342"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_43"
/notes="Organ: eye; Vector: pOTB7; Site:1; XhoI; Site:2;
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"

Query Match 12.9%; Score 736; DB 13; Length 904;
Best Local Similarity 99.6%; Pred. No. 3.9e-100;
Matches 747; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 2342 CGATGCTGTGATGACGATGACATGACGGTGTGACCGATGAGAGGACCACTGCCA 2401
Db 1 CGATGCTGTGATGACGATGACATGACGGTGTGACCGATGAGAGGACCACTGCCA 60
Qy 2402 GCTCCTCTTCAATCCCGCCAGGCTGACTATGACAAAGGATGAGGTGGGACCGCTGTGA 2461
Db 61 GCTCCTCTTCAATCCCGCCAGGCTGACTATGACAAAGGATGAGGTGGGACCGCTGTGA 120

```

```

Qy 2462 CAACTGCCCTTACGTGCACAAACCTCTCCAGATCGACACAGACAACAACTGGAGGGTGA 2521
Db 121 CAACTGCCCTTACGTGCACAAACCTCTCCAGATCGACACAGACAACAACTGGAGGGTGA 180
Qy 2522 CGCCTGCTCCGTGGACATTTGATGGGACGATGTCTTCAATGAACGAGACAATTTGTCCTTA 2581
Db 181 CGCCTGCTCCGTGGACATTTGATGGGACGATGTCTTCAATGAACGAGACAATTTGTCCTTA 240
Qy 2582 CGTCTACACACTGTACCAGAGGACACGGATGTGACGGTGTGGGGATCACTGTGACAA 2641
Db 241 CGTCTACACACTGTACCAGAGGACACGGATGTGACGGTGTGGGGATCACTGTGACAA 300
Qy 2642 CTGCCCCCTGTGTGCACAAACCTTGACACCGACGATGGAACAATGACCTTTGTTGGGACCA 2701
Db 301 CTGCCCCCTGTGTGCACAAACCTTGACACCGACGATGGAACAATGACCTTTGTTGGGACCA 360
Qy 2702 GTGTGACACAAACGAGGACATAGATGACGACGCGCCACAGAACAAACAGGACAACTGGCC 2761
Db 361 GTGTGACACAAACGAGGACATAGATGACGACGCGCCACAGAACAAACAGGACAACTGGCC 420
Qy 2762 CTACATCTCCAAACGCAACGAGCTGACCATGACAGAGAGCGGCGGACGCTGTGA 2821
Db 421 CTACATCTCCAAACGCAACGAGCTGACCATGACAGAGAGCGGCGGACGCTGTGA 480
Qy 2822 CCTGTATGATGACAAACGATGGCGCTCCCGATGACAGGACAACTGCCGCTTGTCTTCAA 2881
Db 481 CCTGTATGATGACAAACGATGGCGCTCCCGATGACAGGACAACTGCCGCTTGTCTTCAA 540
Qy 2882 CCAGACCAAGGAGGACTTGGACCGGTGATGACGCGGTGATATTTGTAAGATGATTTGA 2941
Db 541 CCAGACCAAGGAGGACTTGGACCGGTGATGACGCGGTGATATTTGTAAGATGATTTGA 600
Qy 2942 CAATGACACATCCCAAGATATTGATGATGTGCTCTGAAACAAATGCGATCAGTGAGAC 3001
Db 601 CAATGACACATCCCAAGATATTGATGATGTGCTCTGAAACAAATGCGATCAGTGAGAC 660
Qy 3002 AGACTTCAGAACTTCCAGATGTCCCTTGGATGCCAAAGGACCA-CCGAAATGATC 3060
Db 661 AGACTTCAGAACTTCCAGATGTCCCTTGGATGCCAAAGGACCA-CCGAAATGATC 720
Qy 3061 CCAACTGGGTGCTTCCGCTCAAGGCAAGG 3090
Db 721 CCAACTGGGTGCTTCCGCTCAAGGCAAGG 750

RESULT 15
BQ618192/c
LOCUS BQ618192
DEFINITION UI-H-FHI-bfe-k-04-0-UI.s1 NCI CGAP FH1 Homo sapiens cDNA clone
UI-H-FHI-bfe-k-04-0-UI.3', mRNA sequence.
ACCESSION BQ618192.1 GI:23284407
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 750)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 637-685, >(CAAA)n#Simple_repeat

```


Klausner, R.D., Collins, F.S., Wagner, L., Shennen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Gale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Worley, K.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Wray, J., Helton, E., Kettelman, M., Madan, A.C., Rodriguez, S., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2388257

1247732

2 (bases 1 to 1938)

Strausberg, R.

Direct Submission

Submitted (15-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: amadansystemsbiology.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Place: 19 Row: k Column: 23

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505594.

Location/Qualifiers

1..1938

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:13616 IMAGE:4281085"

/tissue_type="Brain, Primitive neuroectodermal"

/clone_lib="NIH_MGC_58"

/lab_host="DH10B"

/note="Vector: PDNR-LTB"

1..1938

/gene="SERPINE2"

/note="synonyms: PA12, PLANH2, PA1"

/db_xref="LocusID:5055"

/db_xref="MIM:173390"

80..1327

/codon_start=1

/product="serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 2"

/protein_id="AAH12609.1"

/db_xref="GI:15214956"

/db_xref="LocusID:5055"

/translation="MEDLCVANTLFAALNFKHLAKSPQNLFLSPWSISSTWAVMYGSRGSTDQMAVLQNEVGANAFTMTPEFTSCGFMQIQKSYSPDAILQQAADKIHSSPSRSLASNAFTGNLLESVKNLFGKESAFREYIRLQCKYSSSPQVDFLECAEARKKIYSVKTQTKGINLLPGSVGDRVLI VNAVYFKGWKTPFEKLLNGLYPFRVNSAQRTPVQMWYLRKLNIGVIEDLKAQILLELPVAGDVSPFLILPDEIADVS

TGLELESELTVDKLNKWTSKDKVAEDVEVYVPOFKLEHVELRSLRSMGMDAPN
KGRANFGMSERNDLFLSEVPHQAMVDNEEGTEAAAGTGGVMTGTHGGPQFVADH
PFLFLNHKHTNCILFFGRFSSP"

116..1324

/note="SERPIN; Region: SERine Proteinase Inhibitors"

/db_xref="CDD:smart00093"

ORIGIN

Query Match 99.3%; Score 1894.8; DB 9; Length 1938;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GAGGAGCATTCGCCGTCAGACGACGACCACTCAGAGAATAACACAGAGAACAACCAATGAAA 60
Db 18 GAGGAGCATTCGCCGTCAGACGACGACCACTCAGAGAATAACACAGAGAACAACCAATGAAA 77
Qy 61 CAATGGAGGATCTTTGTGTGGCAAAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGG 120
Db 78 CAATGGAGGATCTTTGTGTGGCAAAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGG 137
Qy 121 CAAAAGCAAGCCCAACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGG 180
Db 138 CAAAAGCAAGCCCAACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGG 197
Qy 181 CCATGGTCTATCGGCTCCAGGGCAGACCCGAGACGACGATGGCCAGGTGCTTCAGT 240
Db 198 CCATGGTCTATCGGCTCCAGGGCAGACCCGAGACGATGGCCAGGTGCTTCAGT 257
Qy 241 TTAATGAAGTGGGAGCCCAATCAGATTACCCCATGACTCCAGAGAACTTTTACAGCTGTG 300
Db 258 TTAATGAAGTGGGAGCCCAATCAGATTACCCCATGACTCCAGAGAACTTTTACAGCTGTG 317
Qy 301 GGTTTCATGACGAGATCCAGAGGGGTAGTTATCTCTGATGCGATTTTGAGGCGACAAGCTG 360
Db 318 GGTTTCATGACGAGATCCAGAGGGGTAGTTATCTCTGATGCGATTTTGAGGCGACAAGCTG 377
Qy 361 CAGATAAATCATTCATCTCTCCGCTCTCTCAGCTCTGCAATCAATGATGATCCACAGGGA 420
Db 378 CAGATAAATCATTCATCTCTCCGCTCTCTCAGCTCTGCAATCAATGATGATCCACAGGGA 437
Qy 421 ATTATTTACTGGAAAGTGTCAATAAGCTGTTTGGTGAGAAAGTCTGCGAGGTTCCGGGAAG 480
Db 438 ATTATTTACTGGAAAGTGTCAATAAGCTGTTTGGTGAGAAAGTCTGCGAGGTTCCGGGAAG 497
Qy 481 AATATTTTCGACTCTGTCAGAAATATTACTCTCAGAACCCCGAGGAGTAGACTTCTCTAG 540
Db 498 AATATTTTCGACTCTGTCAGAAATATTACTCTCAGAACCCCGAGGAGTAGACTTCTCTAG 557
Qy 541 AATGTCGAGAGAGCTAGAGAAAAGATTAATTTCTCGGTCAAGACTCAAAACCAAGGCA 600
Db 558 AATGTCGAGAGAGCTAGAGAAAAGATTAATTTCTCGGTCAAGACTCAAAACCAAGGCA 617
Qy 601 AAATCCCAAACTTGTACTCTGAAGGTTCTGTAGATGGGATACAGGATGGTCTCTGTGA 660
Db 618 AAATCCCAAACTTGTACTCTGAAGGTTCTGTAGATGGGATACAGGATGGTCTCTGTGA 677
Qy 661 ATGCTGTCTACTTCAAGAGAAAGTGGAAATCCATTTCAGAGAAACTAAATGGGCTTT 720
Db 678 ATGCTGTCTACTTCAAGAGAAAGTGGAAATCCATTTCAGAGAAACTAAATGGGCTTT 737
Qy 721 ATCTTTCCGTTGTAACCTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGGTGAAA 780
Db 738 ATCTTTCCGTTGTAACCTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGGTGAAA 797
Qy 781 AGCTAAACATTCGATACATAGAGACCTAAAGGCTCAGATTCTTAGAACTCCCATATGCTG 840
Db 798 AGCTAAACATTCGATACATAGAGACCTAAAGGCTCAGATTCTTAGAACTCCCATATGCTG 857
Qy 841 GAGATGTTAGCATGTTCTTGTGCTCCAGATGAAATTCGCGATGTTCTCCACTGGCTGG 900
Db 858 GAGATGTTAGCATGTTCTTGTGCTCCAGATGAAATTCGCGATGTTCTCCACTGGCTGG 917
Qy 901 AGCTCTGGAAAGTGAATAACCTATGACAAACTCAACAAAGTGGACCAAGCAAGACAAA 960

FEATURES
source

gene

CDS

```
108064.1 GI:589223
VERSION: I08064.1
KEYWORDS:
SOURCE: Unknown.
ORGANISM: Unclassified.
REFERENCE: 1 (bases 1 to 1920)
AUTHORS: Webb, A.C. and Auron, P.B.
TITLE: An arg-serpin human plasminogen activator inhibitor designated PAI-2
JOURNAL: Patent: EP 0278696-A2 1 17-AUG-1988;
FEATURES: Location/Qualifiers
           1..1920
           /organism="unknown"
           /mol_type="unassigned DNA"

ORIGIN
Query Match      98.8%; Score 1886; DB 6; Length 1920;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1900; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 GAGGAGCATTGCCCGTCAGACAGCAACTCAGAGAAATACCCAGAGACACACCAGATTGAAA 60
Db 11 GAGGAGCATTGCCCGTCAGACAGCAACTCAGAGAAATACCCAGAGACACACCAGATTGAAA 70
QY 61 CAATGGAGGATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGG 120
Db 71 CAATGGAGGATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGG 130
QY 121 CAAAGCAAGCCCAACCCAGAGCACTCTTCTCTCCCATGGAGCATCTCGTCCACCATGG 180
Db 131 CAAAGCAAGCCCAACCCAGAGCACTCTTCTCTCCCATGGAGCATCTCGTCCACCATGG 190
QY 181 CCATGGTCTACATGGGCTCCAGGGGACGACCGGAAGACAGATGGCCAAAGTCTTCAGT 240
Db 191 CCATGGTCTACATGGGCTCCAGGGGACGACCGGAAGACAGATGGCCAAAGTCTTCAGT 250
QY 241 TTAATGAAGTGGGAGCCAAATGCAGTTACCCCATCACTCCAGAGAACTTTACAGCTGTG 300
Db 251 TTAATGAAGTGGGAGCCAAATGCAGTTACCCCATCACTCCAGAGAACTTTACAGCTGTG 310
QY 301 GGTTTCATGCGAGCAGATCCAGAGGGTAGTTATCTGTAGTCGATTTTTCAGGGCAAGCTG 360
Db 311 GGTTTCATGCGAGCAGATCCAGAGGGTAGTTATCTGTAGTCGATTTTTCAGGGCAAGCTG 370
QY 361 CAGATAAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGA 420
Db 371 CAGATAAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGA 430
QY 421 ATTATTTACTGGAAAGTGTCAATAGCTGTTTGGTGAAGTCTGCGAGCTTCGGGAAG 480
Db 431 ATTATTTACTGGAAAGTGTCAATAGCTGTTTGGTGAAGTCTGCGAGCTTCGGGAAG 490
QY 481 AATATATTTCCGACTCTGTGAGAAATATTACTCTCAGAACCCAGGCGAGTAGACTTCCTAG 540
Db 491 AATATATTTCCGACTCTGTGAGAAATATTACTCTCAGAACCCAGGCGAGTAGACTTCCTAG 550
QY 541 AATGTGAGAGAGACTAGAAAAGATTAATTCCTGGTCAAGACTCAAAACCAAGGCA 600
Db 551 AATGTGAGAGAGACTAGAAAAGATTAATTCCTGGTCAAGACTCAAAACCAAGGCA 610
QY 601 AAATCCCAAACTCTGTACTGCAAGGTTCTGTAGATGGGATACCCAGGATGGTCTCGTGA 660
Db 611 AAATCCCAAACTCTGTACTGCAAGGTTCTGTAGATGGGATACCCAGGATGGTCTCGTGA 670
QY 661 ATGCTGTCTACTTCCAAAGGAAGTGGAAATCTCCATTTGAGAGAACTAAATGGGCTTT 720
Db 671 ATGCTGTCTACTTCCAAAGGAAGTGGAAATCTCCATTTGAGAGAACTAAATGGGCTTT 730
QY 721 ATCCTTTCCGCTGTAACCTCGGCTCAGCGCACACCTGTACAGATGATGTACTTCGCTGAAA 780
Db 731 ATCCTTTCCGCTGTAACCTCGGCTCAGCGCACACCTGTACAGATGATGTACTTCGCTGAAA 790
QY 781 AGCTAAACATTGGGATACATAGAACCTTAAGGCTCAGATTCTTAGAACTCCCATATGCTG 840
```

```
RESULT 2
I08064
LOCUS: I08064
DEFINITION: Sequence 1 from Patent EP 0278696.
ACCESSION: I08064
PAT 02-DEC-1994
linear
DNA
1920 bp
```


Db 791 AGCTAAACATTTGGATACATAGAGACCTAAAGGCTCAGATTCAGAACTCCCATGCTG 850
|||
841 GAGATGTTAGCATGTTCTTGTGCTTCCAGATGAATTTGCCGATGTGTCCACTGGCTTGG 900
|||
851 GAGATGTTAGCATGTTCTTGTGCTTCCAGATGAATTTGCCGATGTGTCCACTGGCTTGG 910
|||
901 AGCTCTCGAAAGTGAATAAATCTATGACAAATCTCAACAGTGGACCCAGCAAAACAAAA 960
|||
911 AGCTCTCGAAAGTGAATAAATCTATGACAAATCTCAACAGTGGACCCAGCAAAACAAAA 970
|||
961 TGGCTGAGATGAAGTGGAGTATACATACCCAGTTCRAATTAGAGAGCATTTATGAAC 1020
|||
971 TGGCTGAGATGAAGTGGAGTATACATACCCAGTTCRAATTAGAGAGCATTTATGAAC 1030
|||
1021 TCAGATCCATTTCTGAAAGCATGGGATGAGAGGAGCGCTTCAACAGGAGCGGCCCAATT 1080
|||
1031 TCAGATCCATTTCTGAAAGCATGGGATGAGAGGAGCGCTTCAACAGGAGCGGCCCAATT 1090
|||
1081 TCTCAGGATGTCGGAGAGGAATGACCTGTTCTTCTGAAAGTGTCCACCAAGCCCATGG 1140
|||
1091 TCTCAGGATGTCGGAGAGGAATGACCTGTTCTTCTGAAAGTGTCCACCAAGCCCATGG 1150
|||
1141 TGGATGTAATGAGGAGGACACTGAGACGCGCTGCGCAGAGGAGTGTATGACAGGGA 1200
|||
1151 TGGATGTAATGAGGAGGACACTGAGACGCGCTGCGCAGAGGAGTGTATGACAGGGA 1210
|||
1201 GAACCTGGATCGAGGCGCCACAGATTTGTGGCAGATCATCCGTTCTTTCTTATTATGC 1260
|||
1211 GAACCTGGATCGAGGCGCCACAGATTTGTGGCAGATCATCCGTTCTTTCTTATTATGC 1270
|||
1261 ATAAGATACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCCTTAAACCTAAGCGT 1320
|||
1271 ATAAGATACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCCTTAAACCTAAGCGT 1330
|||
1321 GCTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCGCTCAGAAATGCTATTTCAAAT 1380
|||
1331 GCTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCGCTCAGAAATGCTATTTCAAAT 1390
|||
1381 TGCCAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAACTTCTGCTACC 1440
|||
1391 TGCCAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAACTTCTGCTACC 1450
|||
1441 CACTAAATTAACACACAGAAATTAATGACAAATGCTCTTATTAACATGACCAACCTATT 1500
|||
1451 CACTAAATTAACACACAGAAATTAATGACAAATGCTCTTATTAACATGACCAACCTATT 1510
|||
1501 AATCATTTGGTCTTCTAAATGGATCATGCCATTTTAGATTTTCTTACTATCAGTTTA 1560
|||
1511 AATCATTTGGTCTTCTAAATGGATCATGCCATTTTAGATTTTCTTACTATCAGTTTA 1570
|||
1561 TTTTATTAACATTAATTTTACTTTGTTATTTATTTATTTATATAATGTTGAGTTTATA 1620
|||
1571 TTTTATTAACATTAATTTTACTTTGTTATTTATTTATTTATATAATGTTGAGTTTATA 1630
|||
1621 ATTATGCTCAGTCCCTATTATTAATGAGCTAAATTAAGTTATAGAGCAGATGATCTGTTA 1680
|||
1631 ATTATGCTCAGTCCCTATTATTAATGAGCTAAATTAAGTTATAGAGCAGATGATCTGTTA 1690
|||
1681 ATTTCCTATCTAATAATGCTTTAATTTGTTCTCATTAATGAAGAAATAGTAGGATCCCT 1740
|||
1691 ATTTCCTATCTAATAATGCTTTAATTTGTTCTCATTAATGAAGAAATAGTAGGTA-CCCT 1749
|||
1741 CCATGCCCTCTGTAATAATCTGAAATAACATTAAACAAATAGGCAATATATGTTA 1800
|||
1750 CCATGCCCTCTGTAATAATCTGAAATAACATTAAACAAATAGGCAATATATGTTA 1809
|||
1801 TGTGCATTTCTAGAAATACATAATATATATGCTGTAATCTATTAATCAATTGCAAG 1860
|||
1810 TGTGCATTTCTAGAAATACATAATATATATGCTGTAATCTATTAATCAATTGCAAG 1869
|||
1861 TATATAATAATAAATCTGCTTCCAAACACAAATAAATAAATAAATAAATAAATAAATAA 1906
|||

Db 1870 TATATAATAATAAATACCTGCTTCCAAACACAAATAAATAAATAAATAAATAA 1915
|||
RESULT 3
A10352
LOCUS
DEFINITION Artificial gene fore plasminogenactivator-inhibitor type 2 (PAI-2).
ACCESSION A10352
VERSION A10352.1 GI:412084
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
JOURNAL
FEATURES
source
Patent: DE 3722673-A 1 19-JAN-1989;
Location/Qualifiers
1..2424
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
553..1800
/codon_start=1
/transl_table=11
/product="plasminogenactivator-inhibitor type 2 (PAI-2)"
/protein_id="CAA00891.1"
/db_xref="GI:412084"
/translation="MEDLCVANTFLALNLFPHLAKASPTQNLFLSPMSISSTMAVYM
GSRGTEQMAKVLQFNEVGANAATPMTPEFTSCGPMQIQKGSYDAILQQAADK
IHSPFSSAINASTGNLLEBSVNLFGKESASPREYIRLCOKYSESEPOAVDFLE
CAEARKINSWVKTKGKIENLPEGSVDGTRMVLVNAVYFKGWKTFPEKFLNG
LYPFVNSAORTPVOMMYLREKLNIGVIEDLKAQIILELPVAGDVSMFLLPDEIADVS
TGLELLESITDYDKNKTSKDMAEDVEVYIPQKLESHYEILRSLSMGMEADFN
KGRANFSGMSERNDLPLSEVFHQAQVNDVNEEGTEAAAGTGGVMTGRTGHGGPQFVADH
PFLFLMHKHTNCILLPFGFSSP"
Query Match 98.5%; Score 1880; DB 6; Length 2424;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1883; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 19 GACAGCAACTCAGAGAAATAACAGAGAAACAACAGATTTGAACCAATGGAGGATCTTTGTG 78
DB 509 GACAGCAACTCAGAGAAATAACAGAGAAACAACAGATTTGAACCAATGGAGGATCTTTGTG 568
QY 79 TGGCAACACACTCTTTGGCTCAATTTATTCAAGCATCTGGCAAGAGCCAGCCACCC 138
DB 569 TGGCAACACACTCTTTGGCTCAATTTATTCAAGCATCTGGCAAGAGCCAGCCACCC 628
QY 139 AGAAGCTCTTCTCTCCCATGGAGCATCTGTCACCATGGCCATGGCTTACATGGGCT 198
DB 629 AGAAGCTCTTCTCTCCCATGGAGCATCTGTCACCATGGCCATGGCTTACATGGGCT 688
QY 199 CAGAGGCGAGCCAGAGACAGATGGCCAGAGTGTTCAGTTTAAATGAAGTGGAGGCA 258
DB 689 CAGAGGCGAGCCAGAGACAGATGGCCAGAGTGTTCAGTTTAAATGAAGTGGAGGCA 748
QY 259 ATGCAATTACCCCATGACTCCAGAGAACTTTTACAGCTGTGGTTCTATGAGCAGATCC 318
DB 749 ATGCAATTACCCCATGACTCCAGAGAACTTTTACAGCTGTGGTTCTATGAGCAGATCC 808
QY 319 AGAAGGCTAGTTATCTCTGATGCGATTTTGCAGGCACAAAGTGCAGATAAAATCCATTCA 378
DB 809 AGAAGGCTAGTTATCTCTGATGCGATTTTGCAGGCACAAAGTGCAGATAAAATCCATTCA 868
QY 379 CCTTCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGAAATTTATTACTTGGAAAGTG 438
DB 869 CCTTCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGAAATTTATTACTTGGAAAGTG 928
QY 439 TCAATAAGCTGTTTGGTGAAGTCTGCGAGCTTCCGGGAGAAATATATGACCTCTGTC 498
DB 929 TCAATAAGCTGTTTGGTGAAGTCTGCGAGCTTCCGGGAGAAATATATGACCTCTGTC 988
|||

ORIGIN

499	AGAAATATTACTCTCTCAGAACCCAGGCGAGTAGACTTCCTAGATGTGCAGAAAGACTA	558
989	AGAAATATTACTCTCTCAGAAACCCAGGCGAGTAGACTTCCTAGATGTGCAGAAAGACTA	1048
559	GAARAAAGATTAACTCTCGGTCAAGACTCAACACCAAGCGCAAAATCCCAAACTTGTTAC	618
1049	GAARAAAGATTAACTCTCGGTCAAGACTCAACACCAAGCGCAAAATCCCAAACTTGTTAC	1108
619	CTGAAGTTCTGTAGATGGGATACCAGATGGTCTGTGGTAATGTGTCTACTCTCAAG	678
1109	CTGAAGTTCTGTAGATGGGATACCAGATGGTCTGTGGTAATGTGTCTACTCTCAAG	1168
679	GAAGTGGAAAATCCATTTGAGAAGAACTAAATCGGCTTTATCTCTCCGTGTAACCT	738
1169	GAAGTGGAAAATCCATTTGAGAAGAACTAAATCGGCTTTATCTCTCCGTGTAACCT	1228
739	CGGCTCAGCGCACACTGTGTACAGATGATGTACTGTGGTGAAGACTAAACATTTGGATACA	798
1229	CGGCTCAGCGCACACTGTGTACAGATGATGTACTGTGGTGAAGACTAAACATTTGGATACA	1288
799	TAGAAGACCTAAAGGCTCAGATTCTAGAATCCCATATGCTGGAGATGTTAGCATGTTCT	858
1289	TAGAAGACCTAAAGGCTCAGATTCTAGAATCCCATATGCTGGAGATGTTAGCATGTTCT	1348
859	TGTTGCTTCCAGATGAAATTCGCGATGTGTCTACTGGCTTGGAGCTGTCTGGAAAGTGAAA	918
1349	TGTTGCTTCCAGATGAAATTCGCGATGTGTCTACTGGCTTGGAGCTGTCTGGAAAGTGAAA	1408
919	TAACTTATGACAACTCAACAAGTGCACAGCAAGACAAAATGGCTGAAGATGAAAGTTG	978
1409	TAACTTATGACAACTCAACAAGTGCACAGCAAGACAAAATGGCTGAAGATGAAAGTTG	1468
979	AGGTATACATACCCAGTTCAAATTTAAGAGGCAATTTAGAACTCAGATCCATTTCTGAAA	1038
1469	AGGTATACATACCCAGTTCAAATTTAAGAGGCAATTTAGAACTCAGATCCATTTCTGAAA	1528
1039	GCATGGCATGGAGGACGCTTCAAACAAGGACGGCGCAATTTCTCAGGATGTCGGAGA	1098
1529	GCATGGCATGGAGGACGCTTCAAACAAGGACGGCGCAATTTCTCAGGATGTCGGAGA	1588
1099	GGAACTGACTGTTCTTTCTGAAAGTGTTCACAAGGCAATGGTGGATGTGAATGAGGAGG	1158
1589	GGAACTGACTGTTCTTTCTGAAAGTGTTCACAAGGCAATGGTGGATGTGAATGAGGAGG	1648
1159	GCACTGAAGCAGCGCTGGCACAGGAGGTGTATGACAGGAGAACTGACATGAGGAGCC	1218
1649	GCACTGAAGCAGCGCTGGCACAGGAGGTGTATGACAGGAGAACTGACATGAGGAGCC	1708
1219	CACAGTTTGTGGAGATCATCCGTTTCTTTTCTTATTATGATGATAGATTAACCAAGTGCA	1278
1709	CACAGTTTGTGGAGATCATCCGTTTCTTTTCTTATTATGATGATAGATTAACCAAGTGCA	1768
1279	TTTTTATTTTTCGGCAGATTTTGTCAACCTTAAACCTAAAGCGTGCTCTGCAAAAGAT	1338
1769	TTTTTATTTTTCGGCAGATTTTGTCAACCTTAAACCTAAAGCGTGCTCTGCAAAAGAT	1828
1339	TTTTGTAGATGAGCTGTGCGCTCAGAAATTTGCTATTTCAAATTTGCCAAAATTTAGAGAT	1398
1829	TTTTGTAGATGAGCTGTGCGCTCAGAAATTTGCTATTTCAAATTTGCCAAAATTTAGAGAT	1888
1399	GTTTTCTACATATTTCTGCTCTTCTGAACAACTTCTGCTACCCACTAAAATAAAAAACAG	1458
1889	GTTTTCTACATATTTCTGCTCTTCTGAACAACTTCTGCTACCCACTAAAATAAAAAACAG	1948
1459	AAATAAATTAGCAATTTGTCTATTATTAACATGACACCTATTAACTATTGGTCTTCTAA	1518
1949	AAATAAATTAGCAATTTGTCTATTATTAACATGACACCTATTAACTATTGGTCTTCTAA	2008
1519	AATGGGATCATGCCAATTTAGATTTTCTTACTATCATGTTTATTTTATTAACATTTAACTT	1578
2009	AATGGGATCATGCCAATTTAGATTTTCTTACTATCATGTTTATTTTATTAACATTTAACTT	2068

Qy	1579	TTACTTTGTTATTTATTTATATAAATGGTGAGTTTTTAAATTTATGCTCACAGCTCA	1638
Db	2069	TTACTTTGTTATTTATTTATATAAATGGTGAGTTTTTAAATTTATGCTCACAGCTCA	2128
Qy	1639	TTTAAATGTAGCTAAATAAGGTTATAGAAGCAGATGATCTGTTAAATTTCTATCTAATAAT	1698
Db	2129	TTTAAATGTAGCTAAATAAGGTTATAGAAGCAGATGATCTGTTAAATTTCTATCTAATAAT	2188
Qy	1699	GCCTTTAAATGTTCTCATATATGAGAAATAGTAGGTATCCCTCCATGCGCTTCTGTAATA	1758
Db	2189	GCCTTTAAATGTTCTCATATATGAGAAATAGTAGGTATCCCTCCATGCGCTTCTGTAATA	2248
Qy	1759	AATATCTCGAAAAAACAATTAACCAATAGGCAATATATGTTATGTGCAATTTCTAGAAATA	1818
Db	2249	AATATCTCGAAAAAACAATTAACCAATAGGCAATATATGTTATGTGCAATTTCTAGAAATA	2308
Qy	1819	CATAACACATATATATGCTGTATCTTATATTCAAATGCGAGTATATATAATAATAAACCT	1878
Db	2309	CATAACACATATATATGCTGTATCTTATATTCAAATGCGAGTATATATAATAATAAACCT	2368
Qy	1879	GCTTCCAAACAACAAAAATAAAAAAA 1906	
Db	2369	GCTTCCAAACAACAAAAATAAAAAAA 2396	
RESULT 4			
AR380643		1884 bp DNA	linear PAT 18-DEC-2003
LOCUS		Sequence 1188 from patent US 6607879.	
DEFINITION		AR380643	
ACCESSION		AR380643	
VERSION		AR380643.1 GI:40088277	
KEYWORDS			
SOURCE		Unknown.	
ORGANISM		Unclassified.	
REFERENCE		1 (bases 1 to 1884)	
AUTHORS		Cocks, B.G., Stuart, S.G. and Seilhamer, J.J.	
TITLE		Compositions for the detection of blood cell and immunological response gene expression	
JOURNAL		Patent: US 6607879-A 1188 19-AUG-2003;	
FEATURES		Location/Qualifiers	
source		1..1884	
		/organism="unknown"	
		/mol_type="genomic DNA"	
ORIGIN			
Query Match		98.4%; Score 1876.6; DB 6; Length 1884;	
Best Local Similarity		99.8%; Pred. No. 0;	
Matches 1879; Conservative		0; Mismatches 4; Indels 0; Gaps 0;	
Qy	9	TTGCCCGTCAGACAGCAACTCAGAGATTAACAGAGAACACACAGATTGAACACATGGAG	68
Db	2	TTACCCCGTCAGACAGCAACTCAGAGATTAACAGAGAACACACAGATTGAACACATGGAG	61
Qy	69	GATCTTTGTGTGGCAACACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCAAAAGCA	128
Db	62	GATCTTTGTGTGGCAACACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCAAAAGCA	121
Qy	129	AGCCCCACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGCT	188
Db	122	AGCCCCACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGCT	181
Qy	189	TACATGGGCTCCAGGGGCGACCGGAGACAGATGGCCAAAGTGCTTCAGTTTAATGAA	248
Db	182	TACATGGGCTCCAGGGGCGACCGGAGACAGATGGCCAAAGTGCTTCAGTTTAATGAA	241
Qy	249	GTGGGAGCAATGTCAGTTACCCCATGACATCCAGAGACTTTACACAGCTGTGGTTCATG	308
Db	242	GTGGGAGCAATGTCAGTTACCCCATGACATCCAGAGACTTTACACAGCTGTGGTTCATG	301
Qy	309	CAGCAGATCCAGAGGTTAGTTATCTGTATGCGATTTTTCAGGCGCAAGCTCCAGATAAA	368
Db	302	CAGCAGATCCAGAGGTTAGTTATCTGTATGCGATTTTTCAGGCGCAAGCTCCAGATAAA	361

369 ATCCATTCAATCTTCGCTCTCTAGCTCTGCAATCAATGCAATCCACAGGGAATATTATTA 428
362 ATCCATTCAATCTTCGCTCTCTAGCTCTGCAATCAATGCAATCCACAGGGAATATTATTA 421
429 CTGGAAGTGTCAATAAGCTGTTTGGTGAGAAGTCTGCGAGCTTCGCGGAAGAAATATTT 488
422 CTGGAAGTGTCAATAAGCTGTTTGGTGAGAAGTCTGCGAGCTTCGCGGAAGAAATATTT 481
489 CGACTCTGTGCAAAATATTACTCTCAGAACCCAGGAGTACAGTCTCTAGAAATGCA 548
482 CGACTCTGTGCAAAATATTACTCTCAGAACCCAGGAGTACAGTCTCTAGAAATGCA 541
549 GAAGAAGCTAGAAAAGAAATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCA 608
542 GAAGAAGCTAGAAAAGAAATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCA 601
609 AACTTGTACCTGAAGTCTGTAGATGGGATACAGAGTGTCTGGTGAATGCTGTC 668
602 AACTTGTACCTGAAGTCTGTAGATGGGATACAGAGTGTCTGGTGAATGCTGTC 661
669 TACTTCAAAGGAAAGTGGAAACTCCATTTGAGAAGAACTAAATGGGCTTTATCCCTTC 728
662 TACTTCAAAGGAAAGTGGAAACTCCATTTGAGAAGAACTAAATGGGCTTTATCCCTTC 721
729 COTGTAAACTCGGCTCAGCGCACACCTGTACAGATGATGACTGCTGGAAGACTAAAC 788
722 COTGTAAACTCGGCTCAGCGCACACCTGTACAGATGATGACTGCTGGAAGACTAAAC 781
789 ATTGGATACATAGAAGACCTAAAGGCTCAGATCTAGAACTCCCATATGCTGGAGATGTT 848
782 ATTGGATACATAGAAGACCTAAAGGCTCAGATCTAGAACTCCCATATGCTGGAGATGTT 841
849 AGCATGTTCTTGTGTTCCAGATGAAATTCGCGATGTCTCCATGCTGCTGGAGTCTGTG 908
842 AGCATGTTCTTGTGTTCCAGATGAAATTCGCGATGTCTCCATGCTGCTGGAGTCTGTG 901
909 GAAAGTGAATAACCTATGACAAACTCAACAAGTGGACCAAGCAAGCAAAATGGCTGAA 968
902 GAAAGTGAATAACCTATGACAAACTCAACAAGTGGACCAAGCAAGCAAAATGGCTGAA 961
969 GATGAAGTTGAGTATACATACCCAGTTCAAATAGAAAGCAATTAAGTCAAGTCC 1028
962 GATGAAGTTGAGTATACATACCCAGTTCAAATAGAAAGCAATTAAGTCAAGTCC 1021
1029 ATTCTGAAAGCAATGCGCATGAGGAGCGCTTCAACAAGGAGCGGCAATTTCTCAGGG 1088
1022 ATTCTGAAAGCAATGCGCATGAGGAGCGCTTCAACAAGGAGCGGCAATTTCTCAGGG 1081
1089 ATGTCGAGAGGAATGACCTGTTCTTCTGAAGTTCACCAAGCCATGCTGGATGTG 1148
1082 ATGTCGAGAGGAATGACCTGTTCTTCTGAAGTTCACCAAGCCATGCTGGATGTG 1141
1149 AATGAGAGGCGCACTGAAGCAGCGCTGGCACAGGAGTGTATGACAGGAGAACTGGA 1208
1142 AATGAGAGGCGCACTGAAGCAGCGCTGGCACAGGAGTGTATGACAGGAGAACTGGA 1201
1209 CATGAGGCGCCACAGTTTGGGAGATCATCGCTTCTTTTCTATATGATAAGATA 1268
1202 CATGAGGCGCCACAGTTTGGGAGATCATCGCTTCTTTTCTATATGATAAGATA 1261
1269 ACCAAGTGCAATTTATTTTTCGGCAGATTTTTCCTCACCTCAAACTAAGCGTCTGCTTC 1328
1262 ACCAAGTGCAATTTATTTTTCGGCAGATTTTTCCTCACCTCAAACTAAGCGTCTGCTTC 1321
1329 TGCAAAAGATTTTGTAGATGAGTGTGCTCAGATTCGATTTTCAAAATGCGCAAA 1388
1322 TGCAAAAGATTTTGTAGATGAGTGTGCTCAGATTCGATTTTCAAAATGCGCAAA 1381
1389 ATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAATCTGCTACCCACTAAAT 1448
1382 ATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAATCTGCTACCCACTAAAT 1441

QY 1449 AABAACACAGAAATAATTAGCAATTTGCTATTATAAATGCAATCAACCCCTATTATCAATTT 1508
Db 1442 AABAACACAGAAATAATTAGCAATTTGCTATTATAAATGCAATCAACCCCTATTATCAATTT 1501
QY 1509 GGTCTTTAAATAAGGATCATGCCCATTTAGATTTTCTTACTATCATGATTTATTTATA 1568
Db 1502 GGTCTTTAAATAAGGATCATGCCCATTTAGATTTTCTTACTATCATGATTTATTTATA 1561
QY 1569 ACATTAACTTTTACTTTTGTATTATTATTATTAATAAGTGTGAGTTTAAATTTATTCG 1628
Db 1562 ACATTAACTTTTACTTTTGTATTATTATTATTAATAAGTGTGAGTTTAAATTTATTCG 1621
QY 1629 TCATGCTTATTATTTAGTCTTAATAAGTTATAGAGAGATGATCTGTTAAATTTCTTA 1688
Db 1622 TCATGCTTATTATTTAGTCTTAATAAGTTATAGAGAGATGATCTGTTAAATTTCTTA 1681
QY 1689 TCTAATAAATGCTTTAAATTTGTTCTCATATGAAGAAATAGTAGGTATCCCTCCATGCC 1748
Db 1682 TCTAATAAATGCTTTAAATTTGTTCTCATATGAAGAAATAGTAGGTATCCCTCCATGCC 1741
QY 1749 TCTGTATAATAATCTGGAATAACATTAACAATAGCAATATATGTTATGTCGATT 1808
Db 1742 TCTGTATAATAATCTGGAATAACATTAACAATAGCAATATATGTTATGTCGATT 1801
QY 1809 TCTAGAATAACATACACATATATGTTCTGTTATCTTATTTCAATTCGAAGTATATAAT 1869
Db 1802 TCTAGAATAACATACACATATATGTTCTGTTATCTTATTTCAATTCGAAGTATATAAT 1861
QY 1869 AAATAAACTGCTTCCAAACAAC 1891
Db 1862 AAATAAACTGCTTCCAAACAAC 1884

RESULT 5

HUMPA12 1884 bp mRNA linear PRI 27-APR-1993
Human placental plasminogen activator inhibitor mRNA, complete cds.

LOCUS J02685.1 GI:189544
DEFINITION plasminogen activator inhibitor; serine protease inhibitor.
ACCESSION J02685.1
VERSION
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PubMed
COMMENT

Ye, R.D., Wun, T.C. and Sadler, J.E.
CDNA cloning and expression in Escherichia coli of a plasminogen
activator inhibitor from human placenta
J. Biol. Chem. 262 (8), 3718-3725 (1987)

Original source text: Human placenta, cDNA to mRNA, clone
lambda-PAI-75.1.
Draft entry and computer-readable sequence for [1] kindly provided
by R.D.Ye, 22-JAN-1987.

FEATURES
Location/Qualifiers
1..1884
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
<1..1884
/product="PAI mRNA"
56..1303
/notes="plasminogen activator inhibitor"

mRNA
CDS
protein_id="AAA36413.1"
/translation="MEDLCVANTLPAFLNFKHAKASPTONLFLSPWISSTMMVMYV
GKSGTDMQAKVLQNEVGANAVTPTNFTSCGEMQIQKQSYPAIILQAOAADK
IHSPKLSAINASTGYLLESVNKLFGKSPREYELCKQKYSVSEPAQVDFLE
CAEARKNSWVKITQTKIPNLLPEGSVDGTRWLVNVAIFYGKWKTPFEKING
LYPFRVNSQRTQVMYLFKLNIGYIDKLQIIEPLPYAGDVSMFLLLPEIDAVS
TGLELLESETTYDKLNKNTSKDMAEDEVYIIPQFLBEHYELNLSIRSMGWDAPN

KGRANPSGMSERNDLFLSEVPHQAMVDNEEGTEAAAGTGVMGRTGHGGPQFVADH		355 bp upstream of PstI site.	
Query Match	98.4%;	Score 1876.6;	DB 9; Length 1884;
Best Local Similarity	99.8%;	Pred. No. 0;	
Mismatches	1879;	Conservative.	0; Mismatches 4; Indels 0; Gaps 0;
1	TTGCGCGT	CAGACGAACT	CAGAGATAAC
2	TTACCGCGT	CAGACGAACT	CAGAGATAAC
69	GATCTTTGT	GTGGCAACAC	ACTCTTTTCCCTC
62	GATCTTTGT	GTGGCAACAC	ACTCTTTTCCCTC
129	AGCCCAACC	CAACCTTCTC	CTCCCATGGAGCAT
122	AGCCCAACC	CAACCTTCTC	CTCCCATGGAGCAT
189	TACATGGCT	CCAGGGGAG	CACGACCAAGCAT
182	TACATGGCT	CCAGGGGAG	CACGACCAAGCAT
249	GTGGAGCCA	ATGAGTTAC	CCCAATGACTCC
242	GTGGAGCCA	ATGAGTTAC	CCCAATGACTCC
309	CAGCAGAT	CCAGAGGGT	AGTTATCCTGAT
302	CAGCAGAT	CCAGAGGGT	AGTTATCCTGAT
369	ATCCATATC	CTTCGGTCT	CTGAGCTCGA
362	ATCCATATC	CTTCGGTCT	CTGAGCTCGA
429	CTGAAAGTG	CAATAAGCT	GTTCGGTGAAG
422	CTGAAAGTG	CAATAAGCT	GTTCGGTGAAG
489	CGACTCTGT	CAGAAATAT	CTCCTCAGAA
482	CGACTCTGT	CAGAAATAT	CTCCTCAGAA
549	GAGAAAGTG	AGAAAGAT	TAATTCCTGGG
542	GAGAAAGTG	AGAAAGAT	TAATTCCTGGG
609	AACCTGTAC	CTGAAAGTT	CTGTAGATGG
602	AACCTGTAC	CTGAAAGTT	CTGTAGATGG
669	TACTTCAAG	GAAGTGA	AACTCCATTT
662	TACTTCAAG	GAAGTGA	AACTCCATTT
729	CGTGTAACT	CTGGCTCAG	CGGCACTGTA
722	CGTGTAACT	CTGGCTCAG	CGGCACTGTA
789	ATTGGATAC	ATAGAGACC	TAAAGGCTCA
782	ATTGGATAC	ATAGAGACC	TAAAGGCTCA
849	AGCATGTT	CTTGTGCT	CCAGATGAA
842	AGCATGTT	CTTGTGCT	CCAGATGAA
909	GAAAGTGA	ATTAACCT	TATGACAACT
902	GAAAGTGA	ATTAACCT	TATGACAACT
969	GATGAAGT	GTGAGGTAT	ACATACCCCA

962	GATGAAGTT	CAGGTATAC	ATACATCCCA
1029	ATTCTGAAA	AGCATGGG	CATGGAGAGC
1022	ATTCTGAAA	AGCATGGG	CATGGAGAGC
1089	ATGTCGAG	AGGAATAC	CTGTTTCTCT
1082	ATGTCGAG	AGGAATAC	CTGTTTCTCT
1149	AATGAGAG	GAGCACTG	GAAGCGGCTG
1142	AATGAGAG	GAGCACTG	GAAGCGGCTG
1209	CATGAGG	CCCAAGT	TTTGGGAGAT
1202	CATGAGG	CCCAAGT	TTTGGGAGAT
1269	ACCAAGT	GCATTTAT	TTTTCGGGAG
1262	ACCAAGT	GCATTTAT	TTTTCGGGAG
1329	TGCAAAA	GATTTTGT	AGATGAGCTG
1322	TGCAAAA	GATTTTGT	AGATGAGCTG
1389	ATTTAGAG	TGTTTCTA	CATATTTCTG
1382	ATTTAGAG	TGTTTCTA	CATATTTCTG
1449	AAAAACA	CAGAAATA	ATTAGACAA
1442	AAAAACA	CAGAAATA	ATTAGACAA
1509	GGTCTTCA	AAAAATGG	ATCATGCCC
1502	GGTCTTCA	AAAAATGG	ATCATGCCC
1569	ACATTAAC	TTTACTTT	TATTTATAT
1562	ACATTAAC	TTTACTTT	TATTTATAT
1629	TCACTGCT	ATTTAATG	TATAGAGAT
1622	TCACTGCT	ATTTAATG	TATAGAGAT
1689	TCTAATAA	TGCTTTAT	TGTTCTCATA
1682	TCTAATAA	TGCTTTAT	TGTTCTCATA
1749	TTCTGTA	ATAAATAT	CTGGAAAA
1742	TTCTGTA	ATAAATAT	CTGGAAAA
1809	TCTAGA	ATAACATA	TATATATG
1802	TCTAGA	ATAACATA	TATATATG
1869	AAATAAAC	CTGCTTCA	AAACAAAC
1862	AAATAAAC	CTGCTTCA	AAACAAAC

RESULT 6
 HSPAI2R
 LOCUS HSPAI2R 1900 bp mRNA linear PRI 31-MAR-1995
 DEFINITION Human mRNA for Arg-Serpin (plasminogen activator-inhibitor 2,
 PAI-2).
 ACCESSION Y00630
 VERSION Y00630.1 GI:35267
 KEYWORDS anti-urokinase; plasminogen activator-inhibitor type 2.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Webb, A.C., Collins, K.L., Snyder, S.E., Alexander, S.J.,
Rosenwasser, L.J., Eddy, R.L., Shows, T.B., and Aueron, P.E.
Hunan monocyte Arg-Serpin cDNA: Sequence, chromosomal assignment,
and homology to plasminogen activator-inhibitor
J. Exp. Med. 166 (1), 77-94 (1987)

JOURNAL MEDLINE
PUBMED 3496414
REFERENCE 2 (bases 1 to 1900)
AUTHORS Webb, A.C.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1987) Andrew C. Webb, Department of Biological
Sciences, Wellesley College, Wellesley, MA 02181, USA
*source=LPS-stimulated monocytes; clone=pcD-1214
COMMENT PAI-2 is a member CC of the serine protease inhibitor (serpin)
superfamily. It inhibits urokinase-type plasminogen activator. The
monocyte derived PAI-2 is distinct from the endothelial
cell-derived PAI-1.

FEATURES
source 1..1900
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
73..1320
/note="unnamed protein product; PAI-2 precursor (AA -22 to
393)"
1..1900
/codon_start=1
/protein_id="CAA68666.1"
/db_xref="GI:35268"
/db_xref="GOA:P05120"
/db_xref="SWISS-PROT:P05120"
/translation="MEDICVANTILFALNFKHLKASPTONLFLSPHSISTMAYVM
GSRGSTDQAVLQNEVGANAVTPTNPENFTSCGFMOQLKQSYDAILQQAADK
IHSFSLASNAKLVSGNYLLESVDFGFKSASFREEYRLQOKYSSSEPAVDLE
CAEARKINSWYKTTGKINLLPEGSVDGTRMVLVNAVYKWKTFPEKKLNG
LYPFRVNSQRTYQVMYKREKLNIYIEDLKQILELPYAGDVSMFLLLPDEIADVS
TGLELESITLYDKLANKWTSKDMAEDEVYTPQKLEHYELRSLILRSMGMDAFN
KGRNPSGMSERNDLFLSEVFOHMDVNEEGTEAAAGTGGVMTGRTGHGGPQFVADH
PLFLIMHKITNCILFGRFSSP"
73..138
/notes="signal peptide (AA -22 to -1)"
139..1317
/product="PAI-2 (AA 1 - 393)"
295..297
/note="N-glycosylation site"
415..417
/note="N-glycosylation site"
1087..1089
/note="N-glycosylation site"
1210..1212
/note="active site of PAI-2"
1568..1577
/note="AT-rich sequence (pot. RNase target)"
1597..1609
/note="AT-rich sequence (pot. RNase target)"
1647..1703
/note="AT-rich sequence (pot. RNase target)"

sig_peptide
mat_peptide
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature

Query Match 98.1%; Score 1872.6; DB 9; Length 1900;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1886; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 GAGGACATTGCCGTGAGAGAGCACTCAGAGATACAGAGAACCAACGATTGAAA 60
DB 11 GAGGACATTGCCGTGAGAGAGCACTCAGAGATACAGAGAACCAACGATTGAAA 70
QY 61 CAATGGAGGATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATCAAGCATCTGG 120
DB 71 CAATGGAGGATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGG 130

QY 121 CAAAGCAGAGCCGCCAGAGAACCTTCTTCTCTCCCATGGAGCATCTCGTCCACCATGG 180
DB 131 CAAAGCAGAGCCGCCAGAGAACCTTCTTCTCTCCCATGGAGCATCTCGTCCACCATGG 190
QY 181 CCATGGTCTACATGGGTCCAGGGGAGCAGACCCAGAGACAGAGATGGCCAAAGTGCTTCAGT 240
DB 191 CCATGGTCTACATGGGTCCAGGGGAGCAGACCCAGAGACAGATGGCCAAAGTGCTTCAGT 250
QY 241 TTAATGAAGTGGAGGCAATGAGTACCCCAATGACTCCAGAGAACTTTACCAAGCTGTG 300
DB 251 TTAATGAAGTGGAGGCAATGAGTACCCCAATGACTCCAGAGAACTTTACCAAGCTGTG 310
QY 301 GGTTCATGAGCAGATCCAGAGAGGTAGTTATCTCTGATGCGATTTTCAGGCAACAAGTG 360
DB 311 GGTTCATGAGCAGATCCAGAGAGGTAGTTATCTCTGATGCGATTTTCAGGCAACAAGTG 370
QY 361 CAGATAAATCCATCATCTCTCCGCTCTCTAGCTCTGCAATCAATGCAATCCACAGGGA 420
DB 371 CAGATAAATCCATCATCTCTCCGCTCTCTAGCTCTGCAATCAATGCAATCCACAGGGA 430
QY 421 ATTATTTACTGAAAAGTGTCAATAAGCTGTTGGTGAGAAAGTCTGCGAGCTTCCGGGAAG 480
DB 431 ATTATTTACTGAAAAGTGTCAATAAGCTGTTGGTGAGAAAGTCTGCGAGCTTCCGGGAAG 490
QY 481 ATATATTCGACTCTGTGAGAAATATTTACTCTCAGAACCCAGCAGTAGACTTCCTTAG 540
DB 491 ATATATTCGACTCTGTGAGAAATATTTACTCTCAGAACCCAGCAGTAGACTTCCTTAG 550
QY 541 AATGTGAGAGAGCTAGAAAAGATTAATTTCTGGGTCAAGACTCAAAACCAAGGCA 600
DB 551 AATGTGAGAGAGCTAGAAAAGATTAATTTCTGGGTCAAGACTCAAAACCAAGGCA 610
QY 601 AATCCCAACTGTTTACCTGAGAGGTTCTGTAGATGGGGATACAGAGATGCTCTCTGGTGA 660
DB 611 AATCCCAACTGTTTACCTGAGAGGTTCTGTAGATGGGGATACAGAGATGCTCTCTGGTGA 670
QY 661 ATGCTGTCTACTTCAAGGAAAGTGAATACTCCATTTGAGAGAACTAATATGGCTTT 720
DB 671 ATGCTGTCTACTTCAAGGAAAGTGAATACTCCATTTGAGAGAACTAATATGGCTTT 730
QY 721 ATCTCTTCCGTGTAAACTCGGCTCAGCGCACACTGTACAGATGATGTACTTGGCTGAAA 780
DB 731 ATCTCTTCCGTGTAAACTCGGCTCAGCGCACACTGTACAGATGATGTACTTGGCTGAAA 790
QY 781 AGCTAAACATTGGATACATAGAGACCTTAAGGCTCAGATCTTGAAGTCCCATATGCTG 840
DB 791 AGCTAAACATTGGATACATAGAGACCTTAAGGCTCAGATCTTGAAGTCCCATATGCTG 850
QY 841 GAGATGTTAGCATGTTCTTGTGTTCTTCCAGATGAAAATTTGCCGATGTGTCCACTGG 900
DB 851 GAGATGTTAGCATGTTCTTGTGTTCTTCCAGATGAAAATTTGCCGATGTGTCCACTGG 910
QY 901 AGCTGCTGAAAAGTGAATTAACCTATGACAACTCAACAGTGGACCCAGCAAGCAAAA 960
DB 911 AGCTGCTGAAAAGTGAATTAACCTATGACAACTCAACAGTGGACCCAGCAAGCAAAA 970
QY 961 TGGCTGAAAGTGAAGTTGAGGTATACATACCCAGTTCAAAATTGAAGAGCATTTGAAC 1020
DB 971 TGGCTGAAAGTGAAGTTGAGGTATACATACCCAGTTCAAAATTGAAGAGCATTTGAAC 1030
QY 1021 TCAGATCCATTTGAAAAGCATGGGATGAGGACGCTTCAACAAAGGAGCGGCAATTT 1080
DB 1031 TCAGATCCATTTGAAAAGCATGGGATGAGGACGCTTCAACAAAGGAGCGGCAATTT 1090
QY 1081 TCTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTTCTGAAAGTGTTCACCAAGGCATGG 1140
DB 1091 TCTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTTCTGAAAGTGTTCACCAAGGCATGG 1150
QY 1141 TGGATGTGAATGAGGAGGACCTGAACGCGCTGGCAGAGGAGGTGTATGACAGGGA 1200
DB 1151 TGGATGTGAATGAGGAGGACCTGAACGCGCTGGCAGAGGAGGTGTATGACAGGGA 1210

```

1201 GAACGGACATGAGGCGCCACAGATTGTGGCAGATCATCCGTTCTTTTCTTATATGC 1260
1211 GAACGGACATGAGGCGCCACAGATTGTGGCAGATCATCCCTTTCTTTCTTATATGC 1270
1261 ATAAGTAACCAAGTGCATTTATTTTTCGGCAGATTTTGTCTCACCTAACTAAGCGT 1320
1271 ATAAGTAACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCTAACTAAGCGT 1330
1321 GCTGCTCTCCAAAGATTTTGTAGAGTGTGCTGCTCAGATGCTATTTCAAT 1380
1331 GCTGCTCTCCAAAGATTTTGTAGAGTGTGCTGCTCAGATGCTATTTCAAT 1390
1381 TGCCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGCAACAATTTCTGTACC 1440
1391 TGCCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGCAACAATTTCTGTACC 1450
1441 CACTAATAAACAACAGATTAATAGACATTTCTATATATACATGACACACCTATT 1500
1451 CACTAATAAACAACAGATTAATAGACATTTCTATATATACATGACACACCTATT 1510
1501 AATCATTGCTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCATGTTA 1560
1511 AATCATTGCTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCATGTTA 1570
1561 TTTTATAACATTAACCTTTTACTTTGTTATTTATTTATTTATATATGATGTTTAA 1620
1571 TTTTATAACATTAACCTTTTACTTTGTTATTTATTTATTTATATGATGTTTAA 1630
1621 ATTATTGCTCCTGCTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCATGTTA 1680
1631 ATTATTGCTCCTGCTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCATGTTA 1690
1681 ATTTCCTATCTAATAATGCTTTTAAATGTTCTCATATGATGATTAAGTATGATGTTA 1740
1691 ATTTCCTATCTAATAATGCTTTTAAATGTTCTCATATGATGATTAAGTATGATGTTA 1749
1741 CCATGCTCTCTGTAATAATATCTGTAATAAATCTTAAACATTAAGGCAATATATGTTA 1800
1750 CCATGCTCTCTGTAATAATATCTGTAATAAATCTTAAACATTAAGGCAATATATGTTA 1809
1801 TGTGCTTTCTAGAAATACATACATATATGCTGTATCTTATATTTCAATTTGCAAG 1860
1810 TGTGCTTTCTAGAAATACATACATATATGCTGTATCTTATATTTCAATTTGCAAG 1869
1861 TATATAATAATAAACCTGCTTCCAAACAAC 1891
1870 TATATAATAATAAACCTGCTTCCAAACAAC 1900

RESULT 7
HUMPAI2B 1880 bp mRNA linear PRI 07-JAN-1995
LOCUS Human plasminogen activator inhibitor 2 (PAI-2) mRNA, complete cds.
DEFINITION M18082
ACCESSION M18082.1 GI:189562
VERSION plasminogen activator inhibitor 2.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1880)
AUTHORS Schleuning,W.D., Medcalf,R.L., Hession,C., Rothenbuhler,R., Shaw,A.
and Kruithof,E.K.
TITLE plasminogen activator inhibitor 2: regulation of gene transcription
during phorbol ester-mediated differentiation of U-937 human
histiocyctic lymphoma cells
Mol. Cell. Biol. 7 (12), 4564-4567 (1987)
JOURNAL 88142852
MEDLINE 3325828
PUBMED
COMMENT Original source text: Human histiocyctic lymphoma cells, cDNA to
mRNA, clone pPAIJ.7.
FEATURES Location/Qualifiers
source 1..1880

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="7q21.3-q22.1"
1..1880
/gene="PAI1"
CDS
32..1279
/gene="PAI1"
/notes="plasminogen activator inhibitor type 2 precursor"
/codon_start=1
/protein_id="AAA60006.1"
/db_xref="GI:189563"
/translation="MEDLCVANTLFLALNFKHLAKASPTQNLFLSPWSISSTWAMVYM
GSGSTDDQMAKVLQFNEVGANAVTPTPEFTSCFMQOIQKGSYPDAILOQAADK
IHSGFRSLGAINASTGDIYLLSEVKNLPEGSASPREYIRLCQKYSPPQAVDFLE
CASEARKINSWKTQTKGIPNLLPEGSVDGTRMVLNNAVTFKQKWTTPFKKNG
LYEPVNSAORTPVOMAVLREKNGIYIEDLKAQILELPYAGDVSMFLLLPDIADVS
TGLLELSEITYDKLNWTSKDMAEDREYVIRPKLBEHYELSLILSMGMEDAFN
KGRANFGSMERNDFLSEVHQAVDVNEEGTEAAGTGGVMTGRTGGGPFQVADH
PFLFLIMHKITKCLIFFGRFCSP"
32..97
/gene="PAI1"
/notes="plasminogen activator inhibitor type 2 signal
peptide"
98..1276
/gene="PAI1"
/product="plasminogen activator inhibitor type 2 protein"

```

ORIGIN

```

Query Match 96.9%; Score 1849.2; DB 9; Length 1880;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1865; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 34 AATAACAGAGAACCAACAGATTGAAACAATGGAGATCTTTGTGGCAAAACACACTCT 93
Db 3 ATCCCCAGAGAACCAACAGATTGAAACAATGGAGATCTTTGTGGCAAAACACACTCT 62
QY 94 TTGCCCTCAATTTTATCAAGCATCTGGCAAAAGCAAGCCCCACCCAGAACCTTCTCT 153
Db 63 TTGCCCTCAATTTTATCAAGCATCTGGCAAAAGCAAGCCCCACCCAGAACCTTCTCT 122
QY 154 CCCCATGAGCATCTCGTCCACCATGGCCATGCTTACATGGCTCCAGGGGAGACCG 213
Db 123 CCCCATGAGCATCTCGTCCACCATGGCCATGCTTACATGGCTCCAGGGGAGACCG 182
QY 214 AAGACAGATGGCCAAAGTGTCTTCAAGTTTAAATGAAGTGGAGCCCAATGCAGTTACCCCA 273
Db 183 AAGACAGATGGCCAAAGTGTCTTCAAGTTTAAATGAAGTGGAGCCCAATGCAGTTACCCCA 242
QY 274 TGACTCCAGAACTTTTACAGCTGTGGTTTCATGCAGCATCCAGAGGGTAGTTATC 333
Db 243 TGACTCCAGAACTTTTACAGCTGTGGTTTCATGCAGCATCCAGAGGGTAGTTATC 302
QY 334 CTGATGCGATTTTCAGGCAACAAGTGCAGATAAAATCCATTCATCTCCGCTCTCTCA 393
Db 303 CTGATGCGATTTTCAGGCAACAAGTGCAGATAAAATCCATTCATCTCCGCTCTCTCA 362
QY 394 GCTCTGCAATCAATGCATCCACAGGGAATTTTACTGAAAAGTGTCAATAGCTGTTG 453
Db 363 GCTCTGCAATCAATGCATCCACAGGGAATTTTACTGAAAAGTGTCAATAGCTGTTG 422
QY 454 GTGAGAGTCTGCGAGCTTCGGGAAGATATATTTCGACTGTGTCAAAATATATCTCT 513
Db 423 GTGAGAGTCTGCGAGCTTCGGGAAGATATATTTCGACTGTGTCAAAATATATCTCT 482
QY 514 CAGAACCCCGAGCTAGACTTCTTAGAATGTGCAAGAAAGCTAGAAAAAGATTAATT 573
Db 483 CAGAACCCCGAGCTAGACTTCTTAGAATGTGCAAGAAAGCTAGAAAAAGATTAATT 542
QY 574 CTTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTTTGTTACTGAGGTTCTGTAG 633
Db 543 CTTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTTTGTTACTGAGGTTCTGTAG 602

```

634	ATGGGGATACCAGGATGGTCTGGTGAATGCTGTCTACTCAAGAGAAAGTGGAAAACTC	693
Qy		
603	ATGGGGATACCAGGATGGTCTGGTGAATGCTGTCTACTCAAGAGAAAGTGGAAAACTC	662
Db		
694	CATTTGAGAAGAAACTAAATGGGCTTTATCCTTTCCGTGTAAACTCGGCTCAGGCGCACAC	753
Qy		
663	CATTTGAGAAGAAACTAAATGGGCTTTATCCTTTCCGTGTAAACTCGGCTCAGGCGCACAC	722
Db		
754	CTGTACAGATGATGTACTTGGTGTGAAAAGCTAAACATTTGNAATACATAGAGACCTAAAGG	813
Qy		
723	CTGTACAGATGATGTACTTGGTGTGAAAAGCTAAACATTTGNAATACATAGAGACCTAAAGG	782
Db		
814	CTCAGATTCTAGAACCTCCCATATGCTTGGAGATGTTAGCATGTTCTTGTGTGTTCCAGATG	873
Qy		
783	CTCAGATTCTAGAACCTCCCATATGCTTGGAGATGTTAGCATGTTCTTGTGTGTTCCAGATG	842
Db		
874	AAATTTGCCGATGTGCCATCTGGCTTGGAGCTGCTGGAAAGTGAATACCTATGACAAAC	933
Qy		
843	AAATTTGCCGATGTGCCATCTGGCTTGGAGCTGCTGGAAAGTGAATACCTATGACAAAC	902
Db		
934	TCAAACAAGTGGACCAAGCAAAAGACAAAATGGCTGAAGATGAAGTTGAGGTATACATACCCC	993
Qy		
903	TCAAACAAGTGGACCAAGCAAAAGACAAAATGGCTGAAGATGAAGTTGAGGTATACATACCCC	962
Db		
994	AGTTTCAAATTAGAAGAGCATTTATGMACTCAGATCCATTCTGAAAGCATGGCGATGGAGG	1053
Qy		
963	AGTTTCAAATTAGAAGAGCATTTATGMACTCAGATCCATTCTGAAAGCATGGCGATGGAGG	1022
Db		
1054	ACGCCCTTCAACAAGGACGGCGCCAAATTTCTCAGGGATGTGCGAGAGGAATGACCTGTTTC	1113
Qy		
1023	ACGCCCTTCAACAAGGACGGCGCCAAATTTCTCAGGGATGTGCGAGAGGAATGACCTGTTTC	1082
Db		
1114	TTTCTGAGTGTTCACACAGCCATGCTGGATGTAATGCGAGGGCACTGAAGCAGCGG	1173
Qy		
1083	TTTCTGAGTGTTCACACAGCCATGCTGGATGTAATGCGAGGGCACTGAAGCAGCGG	1142
Db		
1174	CTGGCACAGGAGGTGTTATGACAGGGGAACTGCAATGCGAGGGCCACAGTTTGTGGCAG	1233
Qy		
1143	CTGGCACAGGAGGTGTTATGACAGGGGAACTGCAATGCGAGGGCCACAGTTTGTGGCAG	1202
Db		
1234	ATCATTCGGTTCTTTTCTTTATATGATGAATGAATGCGAGGGCACTGAAGTTCGGCA	1293
Qy		
1203	ATCATTCGGTTCTTTTCTTTATATGATGAATGAATGCGAGGGCACTGAAGTTCGGCA	1262
Db		
1294	GAATTTGCTCACCTTAAACCTAAGCGTCTGCTTCTGCAAAAGATTTTGTGTAGTAGCT	1353
Qy		
1263	GAATTTGCTCACCTTAAACCTAAGCGTCTGCTTCTGCAAAAGATTTTGTGTAGTAGCT	1322
Db		
1354	GTGTGCTCAGAAATGCTATTTTCAAAATGCCCCAAATTTAGAGATGTTTTCTACATATTT	1413
Qy		
1323	GTGTGCTCAGAAATGCTATTTTCAAAATGCCCCAAATTTAGAGATGTTTTCTACATATTT	1382
Db		
1414	CTGCTCTTCTGAACAACTTCTGCTACCGCATTAATATAAAACACAGAAATAATAGACAAT	1473
Qy		
1383	CTGCTCTTCTGAACAACTTCTGCTACCGCATTAATATAAAACACAGAAATAATAGACAAT	1442
Db		
1474	TGFTCTATTATAACATGACAAACCTTATTAACTCAATTTGGTCTTCTAAAAATGGGATCATGCC	1533
Qy		
1443	TGFTCTATTATAACATGACAAACCTTATTAACTCAATTTGGTCTTCTAAAAATGGGATCATGCC	1502
Db		
1534	ATTTAGATTTTCCCTTACTATCAGTTTATTTTATAACATTAACCTTTTACTTTGTATTATA	1593
Qy		
1503	ATTTAGATTTTCCCTTACTATCAGTTTATTTTATAACATTAACCTTTTACTTTGTATTATA	1562
Db		
1594	TTATTTTATATAATGGTGTAGTTTATAATTTATGCTCACTGCTATTTAATGTAGCTAAT	1653
Qy		
1563	TTATTTTATATAATGGTGTAGTTTATAATTTATGCTCACTGCTATTTAATGTAGCTAAT	1621
Db		
1654	AAAGTTATAGAAGCAGATGATCTGTAAATTTCCCTATCTAATAAATGCCCTTTAATGTTCT	1713
Qy		
1622	AAAGTTATAGAAGCAGATGATCTGTAAATTTCCCTATCTAATAAATGCCCTTTAATGTTCT	1681
Db		
1714	CATAATGAAGAAATGAAGTAGGTATCCCTCCATGCCCTTCTGTAAATAAATATCTGGAAAAA	1773
Qy		

Db	1582	CATATGAGAAATAGTAGGTATCCCTCCATGGCCCTTCTATATATATATCTGGAATAA	1741
Qy	1774	CATTAAACAATAGGCAAAATATGTTATGTGCATTTCTAGAAATACATAACATATATA	1833
Db	1742	CATTAAACAATAGGCAAAATATGTTATGTGCATTTCTAGAAATACATAACATATATA	1801
Qy	1834	TGCTGTATCTTATATTCAATTGCAAGTATATATAATAATAAACTGCTTCCAAACAACAA	1893
Db	1802	TGCTGTATCTTATATTCAATTGCAAGTATATATAATAATAAACTGCTTCCAAACAACAA	1861
Qy	1894	AAAAATAAAAAAAG 1907	
Db	1862	AAAAATAAAAAAAG 1875	
RESULT 8			
LOCUS	AR380644	1854 bp DNA linear	PAT 18-DEC-2003
DEFINITION	Sequence 1189 from patent US 6607879.		
ACCESSION	AR380644		
VERSION	AR380644.1	GI:40088278	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1854)		
AUTHORS	Cocks, B. G., Stuart, S. G. and Seilhamer, J. J.		
TITLE	Compositions for the detection of blood cell and immunological response gene expression		
JOURNAL	Patent: US 6607879-A 1189 19-AUG-2003;		
FEATURES	Location/Qualifiers		
source	1..1854		
	/organism="unknown"		
ORIGIN	/mol_type="genomic DNA"		
Query Match	96.8%;	Score 1847.6;	DB 6; Length 1854;
Best Local Similarity	99.8%;	Pred. No. 0;	
Matches 1850;	Conservative 0;	Mismatches 4;	Indels 0; Gaps 0
Qy	15	GTGAGACAGCAACTCAGAGATAACCGAGAGAACCAACAGATTGAAACAAATGAGGATCTT	74
Db	1	GTGAGACAGCAACTCAGAGATAACCGAGAGAACCAACAGATTGAAACAAATGAGGATCTT	60
Qy	75	TGTTGGCAACACACTTTTGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCCCC	134
Db	61	TGTTGGCAACACACTTTTGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCCCC	120
Qy	135	ACCAGAACTCTTCCTCTCCCATGAGCATCTGTCACCATTGGCCATGGTCTACATG	194
Db	121	ACCAGAACTCTTCCTCTCCCATGAGCATCTGTCACCATTGGCCATGGTCTACATG	180
Qy	195	GGCTCCAGGGGCGAGCACCGAAGACCAAGTGGCCAAAGTGCTTCAGTTTAATGAAATGGGA	254
Db	181	GGCTCCAGGGGCGAGCACCGAAGACCAAGTGGCCAAAGTGCTTCAGTTTAATGAAATGGGA	240
Qy	235	GCCATGCAAGTACCCCATGACTCCAGAGAACTTTACCACTGTGGGTTTCATGCAGCAG	314
Db	241	GCCATGCAAGTACCCCATGACTCCAGAGAACTTTACCACTGTGGGTTTCATGCAGCAG	300
Qy	315	ATCCAGAAAGGTAGTTATTCCTGATGCGATTTTGCAAGCACCAAGCTGCAGATAAATCCAT	374
Db	301	ATCCAGAAAGGTAGTTATTCCTGATGCGATTTTGCAAGCACCAAGCTGCAGATAAATCCAT	360
Qy	375	TCATCCTTCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAAATTTATTCTGGAA	434
Db	361	TCATCCTTCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAAATTTATTCTGGAA	420
Qy	435	AGTGTCAATAAGCTGTTTGGTGAGAAGTCTCGAGCTTCCGGGAAGAAATATATTCGATCTC	494
Db	421	AGTGTCAATAAGCTGTTTGGTGAGAAGTCTCGAGCTTCCGGGAAGAAATATATTCGATCTC	480

Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1850; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
2y	15	QTCAGACGAACTCAGAGAAATACAGAGAAACACAGAGATCAAAACAAATGGAGATCTT	74
Db	1	QTCAGACGAACTCAGAGAAATACAGAGAAACACAGAGATCAAAACAAATGGAGATCTT	60
2y	75	TGTGTGGCAACACACTCTTTCCTCCCTCAATTTATTTCAAGCATCTGGCAAAAGCAAGCCCC	134
Db	61	TGTGTGGCAACACACTCTTTCCTCCCTCAATTTATTTCAAGCATCTGGCAAAAGCAAGCCCC	120
Qy	135	ACCCAGAACCTCTTCTCTCCCATCGAGCATCTGTCACCATGGCCATGGCTTACATG	194
Db	121	ACCCAGAACCTCTTCTCTCCCATCGAGCATCTGTCACCATGGCCATGGCTTACATG	180
Qy	195	GGCTCCAGGGGAGACACCAAGACAGATGGCCAAAGTGCTTCAGTTTAAATGAAGTGGGA	254
Db	181	GGCTCCAGGGGAGACACCAAGACAGATGGCCAAAGTGCTTCAGTTTAAATGAAGTGGGA	240
Qy	255	GCCAAATGCAGTTACCCCATGACTCCAGAGAACTTTTACCAAGCTGTGGGTTTCATGCAGAG	314
Db	241	GCCAAATGCAGTTACCCCATGACTCCAGAGAACTTTTACCAAGCTGTGGGTTTCATGCAGAG	300
Qy	315	ATCCAGAGGGTAGTTATCTGTAGTGGATTTTCAGGSCAACAGCTGCAGATAAATCCAT	374
Db	301	ATCCAGAGGGTAGTTATCTGTAGTGGATTTTCAGGSCAACAGCTGCAGATAAATCCAT	360
Qy	375	TCATCTCTCCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGAAATTTTACTGGAA	434
Db	361	TCATCTCTCCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGAAATTTTACTGGAA	420
Qy	435	AGTGTCATTAAGCTGTTGGTGAGAGTCTGCAGCTTCCGGGAAGATATATTCGACTC	494
Db	421	AGTGTCATTAAGCTGTTGGTGAGAGTCTGCAGCTTCCGGGAAGATATATTCGACTC	480
Qy	495	TGTCAGAAATATTACTCTCAGAACCCCGAGAGTAGACTTCCTAGAATGTGCAGAGAA	554
Db	481	TGTCAGAAATATTACTCTCAGAACCCCGAGAGTAGACTTCCTAGAATGTGCAGAGAA	540
Qy	555	GCTAGAAAAGATTAATCTCTGGGTCAAGACTCAACACCAAGGCAAAATCCCAACTTG	614
Db	541	GCTAGAAAAGATTAATCTCTGGGTCAAGACTCAACACCAAGGCAAAATCCCAACTTG	600
Qy	615	TTACCTGAAAGTTCTGTAGATGGGATACAGGATGCTCTGGTGAATGTCTACTTC	674
Db	601	TTACCTGAAAGTTCTGTAGATGGGATACAGGATGCTCTGGTGAATGTCTACTTC	660
Qy	675	AAAGGAAGTGGAAACTCCATTGAGAGAACTAAATGGGCTTTATCTCTCCGTGTA	734
Db	661	AAAGGAAGTGGAAACTCCATTGAGAGAACTAAATGGGCTTTATCTCTCCGTGTA	720
Qy	735	AACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGGTGAAAAGCTAAACATTGGA	794
Db	721	AACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGGTGAAAAGCTAAACATTGGA	780
Qy	795	TACATAGAACCTTAAGGCTCAGATTTCTAGACTCCCATATGCTGGAGATGTTAGCATG	854
Db	781	TACATAGAACCTTAAGGCTCAGATTTCTAGACTCCCATATGCTGGAGATGTTAGCATG	840
Qy	855	TTCTCTGTTGTTCCAGATGAAATTCGCGATGTGTCCTACTGGCTGGAGCTCTGGAAAGT	914
Db	841	TTCTCTGTTGTTCCAGATGAAATTCGCGATGTGTCCTACTGGCTGGAGCTCTGGAAAGT	900
Qy	915	GAAATTAACCTATGACAACTCAACAAGTGGACAGCAAGCAAAATGGCTGAGATGAA	974
Db	901	GAAATTAACCTATGACAACTCAACAAGTGGACAGCAAGCAAAATGGCTGAGATGAA	960
Qy	975	GTTGAGGTATACATACCCCAAGTTCAAAATTAGAGAGCATTTATGAACCTCAGATCCATCTG	1034
Db	961	GTTGAGGTATACATACCCCAAGTTCAAAATTAGAGAGCATTTATGAACCTCAGATCCATCTG	1020
Qy	1035	AAAGCATGGGATGGAGGACGCTTCAACAGGAGCGGCAATTTCTCAGGAGATGTCG	1094

Db	1021	AGAAGCATGGGATGGAGGACGCTTCAACAAAGGACGGGCCCAATTTCTCAGGATGTCG	1080
Qy	1095	GAGAGAAATGACCTTTCTTTCTGAAAGTTCCACCAAGCCATGGTGGATGTGAATGAG	1154
Db	1081	GAGAGAAATGACCTTTCTTTCTGAAAGTTCCACCAAGCCATGGTGGATGTGAATGAG	1140
Qy	1155	GAGGCACTGAAGCAGCCGCTGGCAGAGGAGTGTATGACAGGAGAACTGGACATGGA	1214
Db	1141	GAGGCACTGAAGCAGCCGCTGGCAGAGGAGTGTATGACAGGAGAACTGGACATGGA	1200
Qy	1215	GGCCACAGTTTGTGGCAGCATCCGTTTCTTTTCTTTTCTTTTATGATAGATAACCAAG	1274
Db	1201	GGCCACAGTTTGTGGCAGCATCCGTTTCTTTTCTTTTCTTTTATGATAGATAACCAAG	1260
Qy	1275	TGCAATTTTATTTTGGGAGATTTTGTCTCACCCCTAAAACCTAAGCGTCTGCTCTGCAAA	1334
Db	1261	TGCAATTTTATTTTGGGAGATTTTGTCTCACCCCTAAAACCTAAGCGTCTGCTCTGCAAA	1320
Qy	1335	AGATTTTGTAGATGAGCTGTGCTCAGATTGCTATTTTCAAAATGCCAAATAATTAG	1394
Db	1321	AGATTTTGTAGATGAGCTGTGCTCAGATTGCTATTTTCAAAATGCCAAATAATTAG	1380
Qy	1395	AGATCTTTTCTACATATTTTCTCTCTCTGAAACAACCTTCTGCTACCCACTAAATAAAC	1454
Db	1381	AGATCTTTTCTACATATTTTCTCTCTCTGAAACAACCTTCTGCTACCCACTAAATAAAC	1440
Qy	1455	ACAGAAATAATTAGCAATTTCTATTATTAACATGACACCCCTATTATTAATCATTTGCTT	1514
Db	1441	ACAGAAATAATTAGCAATTTCTATTATTAACATGACACCCCTATTATTAATCATTTGCTT	1500
Qy	1515	CTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTATTTTATTAACATTA	1574
Db	1501	CTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTATTTTATTAACATTA	1560
Qy	1575	ACTTTTACTTTGTTTATTTATTTTATTAATGAGTGGTGGATTTTAAATTTATGCTCCTG	1634
Db	1561	ACTTTTACTTTGTTTATTTATTTTATTAATGAGTGGTGGATTTTAAATTTATGCTCCTG	1620
Qy	1635	CCTATTTATGTAGCTAATAAAGTTATAGAGCAGATGATCTGTAAATTTCTCTATCTAAT	1694
Db	1621	CCTATTTATGTAGCTAATAAAGTTATAGAGCAGATGATCTGTAAATTTCTCTATCTAAT	1680
Qy	1695	AAATGCTTTTATTTGTTCTCAATATGAGATAGTAGGTATCCCTCCATGCCCTCTCTG	1754
Db	1681	AAATGCTTTTATTTGTTCTCAATATGAGATAGTAGGTATCCCTCCATGCCCTCTCTG	1740
Qy	1755	AATAAATATCTGAAAAAACAATTAACCAATAGGCAATATATGTTATGTGCAATTTCTAGA	1814
Db	1741	AATAAATATCTGAAAAAACAATTAACCAATAGGCAATATATGTTATGTGCAATTTCTAGA	1800
Qy	1815	AATACATAACATATATATGCTGTATCTTATTTCAATTTGCAAGTATATAAT	1868
Db	1801	AATACATAACATATATATGCTGTATCTTATTTCAATTTGCAAGTATATAAT	1854

RESULT 10

A02514	A02514	2409 bp	DNA	linear	PAT 24-MAY-1993
LOCUS					
DEFINITION	Synthetic gene for miniactivin.				
ACCESSION	A02514				
VERSION	A02514.1	GI:412314			
KEYWORDS	synthetic construct				
SOURCE	synthetic construct				
ORGANISM	artificial sequences.				
REFERENCE	1 (bases 1 to 2409)				
AUTHORS	Antalis,T.M., Barnes,T.M., Clark,M.A., Devine,P.L., Goss,N.H. and Lehrbach,P.R.				
TITLE	Recombinant product				
JOURNAL	Patent: EP 0238275-A 16 23-SEP-1987;				
FEATURES	BIOTECHNOLOGY AUSTRALIA PTY. LTD.; THE AUSTRALIAN NATIONAL UNIVERSITY				
	Location/Qualifiers				

```

1..2409
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
49..1296
/codon_start=1
/transl_table=11
/product="miniactivin"
/protein_id="CAA00247.1"
/db_xref="GI:412335"
/db_xref="REFSEQ:CAA00247"
/translation="MEDJCVAATFALNLFKHLAKASPTQNLFLSPWSISSTWAMVYM
GSRGTBQMAKVLQFNEGVANAVTTPENTPSTSCFMQOIQKGSYFDALQQAADLE
IHSSRSLTSSANASTGYLLSVNKLFGKESAPREYIRLQCKYISSEPPQAVDLE
CAEARKINSNVKTKGKIENLIPGSDGDTIRMLVNAVYFKGKWTTPFKKDLG
LYPFVNSAQRTPVQMYMLREKLNGYIDLKAQILELPYAGDVSNFKMLLPDIEIADVS
TGLELSEBITVDKLNKMTSKQMAEDEVEVTIPOPKLBEHVELRIILSRMGMEDAFN
KGRANFGMSERNDLFLSVFHCQAVDVNEEGTEAAAGTGGVMTGRTGHGGPQFVADH
PFLFLMHKHTNLCILPFGFSP"

>ORIGIN
      15  GTCAGACACCAACTCAGAGAAATAACAGAGAAACAACAGATTGAAACAATGGAGGATCTT  74
      1  GTCAGACACCAACTCAGAGAAATAACAGAGAAACAACAGATTGAAACAATGGAGGATCTT  60
      75  TGTGTGGCAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCCCC  134
      61  TGTGTGGCAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCCCC  120
      135  ACCAGAACTCTTCCTCCCATGGAGCATCTGTCGTCACCATGGCCATGGTCTACATG  194
      121  ACCAGAACTCTTCCTCCCATGGAGCATCTGTCGTCACCATGGCCATGGTCTACATG  180
      195  GGCTCCAGGGGCGAGCACCGAAGCACAGATGGCCAAAGTGCTTCAGTTTAATGAAGTGGGA  254
      181  GGCTCCAGGGGCGAGCACCGAAGCACAGATGGCCAAAGTGCTTCAGTTTAATGAAGTGGGA  240
      255  GCCAATGCAGTTACCCCATGACTCCAGAGAACTTTACCAAGCTGGGGTTCATGCACAG  314
      241  GCCAATGCAGTTACCCCATGACTCCAGAGAACTTTACCAAGCTGGGGTTCATGCACAG  300
      315  ATCCAGAAGGGTAGTTATCTCGATGCGAATTTGCAGGCACAAGCTGCAGATAAATCCAT  374
      301  ATCCAGAAGGGTAGTTATCTCGATGCGAATTTGCAGGCACAAGCTGCAGATAAATCCAT  360
      375  TCATCTTCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGAATTTATTACTGGAA  434
      361  TCATCTTCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGAATTTATTACTGGAA  420
      435  AGTGTCAATAGCTGTGTTGGTGAAGAGTCTGCGAGCTTCGGGAGAGATATATTGCACTC  494
      421  AGTGTCAATAGCTGTGTTGGTGAAGAGTCTGCGAGCTTCGGGAGAGATATATTGCACTC  480
      495  TGTCAAGAAATATTACTCCTCAGAACCCCGAGCAGTAGACTTCCTAGAATGTGCAGAGAA  554
      481  TGTCAAGAAATATTACTCCTCAGAACCCCGAGCAGTAGACTTCCTAGAATGTGCAGAGAA  540
      555  GCTAGAAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTTG  614
      541  GCTAGAAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTTG  600
      615  TTACTCAAGGTTCTGTAGATGGGGATACCAAGATGTCCTGGTGAATGCTGTACTTTC  674
      601  TTACTCAAGGTTCTGTAGATGGGGATACCAAGATGTCCTGGTGAATGCTGTACTTTC  660
      675  AAAGGAAAGTGGAAAACCTCCATTTGAGAAGAACTAAATGGGCTTTTACTCTTCGTTGA  734
      661  AAAGGAAAGTGGAAAACCTCCATTTGAGAAGAACTAAATGGGCTTTTACTCTTCGTTGA  720

```


2y 1575 ACTTTTACTTGTGTTATTTATTTATATATATGTTGAGTCTTTTAAATTTATGTCACCTG 1634
Db 1561 ACTTTTACTTGTGTTATTTATTTATATATATGTTGAGTCTTTTAAATTTATGTCACCTG 1620
2y 1635 CCTATTTAATGTAGCTAATAAAGTTATAGAAGCAGATGATCTGTTAATTTTCTATCTAAT 1694
Db 1621 CCTATTTAATGTAGCTAATAAAGTTATAGAAGCAGATGATCTGTTAATTTTCTATCTAAT 1680
2y 1695 AAATGCGCTTTAATCTCTCATATGAAGATAAGTAGGTATCCCTCCATGCCCTTCTGT 1754
Db 1681 AAATGCGCTTTAATCTCTCATATGAAGATAAGTAGGTATCCCTCCATGCCCTTCTGT 1740
2y 1755 AATAAATATCTGGAATAAACAANTTAAACAATAGCAAAATATATGTTATGTCATTTCTAGA 1814
Db 1741 AATAAATATCTGGAATAAACAANTTAAACAATAGCAAAATATATGTTATGTCATTTCTAGA 1800
2y 1815 AATACATAACACATATATATGTCGTATCTTATATTCAATTGCAAGTATATATAT 1868
Db 1801 AATACATAACACATATATATGTCGTATCTTATATTCAATTGCAAGTATATATAT 1854

RESULT 12
I12242
LOCUS I12242 2409 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 18 from patent US 5422090.
ACCESSION I12242
VERSION I12242.1 GI:910265
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2409)
AUTHORS Stephens,R.W., Golder,J.P., Antalis,T.M., Barnes,T.M., Clark,M.A., Devine,P.L., Goss,N.H. and Lehrbach,P.R.
TITLE Human PAI-2
JOURNAL Patent: US 5422090-A 18 06-JUN-1995;
FEATURES
Location/Qualifiers
source 1..2409
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 96.8%; Score 1847.6; DB 6; Length 2409;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
15 GTCAGACAGCACTCAGAGAAATACAGAGAAACAACAGATGGAACAATGGAGATCTT 74
1 GTCAGACAGCACTCAGAGAAATACAGAGAAACAACAGATGGAACAATGGAGATCTT 60
75 TGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCCCC 134
61 TGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCCCC 120
135 ACCGAGACCTCTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGCTGTACATG 194
121 ACCGAGACCTCTCTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGCTGTACATG 180
195 GGTCTCCAGGGGAGCAGCAAGCAAGCAAGATGGCAAGGTGTTTCAAGTTAATGAAGTGGGA 254
181 GGTCTCCAGGGGAGCAGCAAGCAAGCAAGATGGCAAGGTGTTTCAAGTTAATGAAGTGGGA 240
255 GCCAATGAGTTACCCCATGACTCCAGAGAACTTTTACCAAGCTGTGGGTTTCATGCAGCAG 314
241 GCCAATGAGTTACCCCATGACTCCAGAGAACTTTTACCAAGCTGTGGGTTTCATGCAGCAG 300
315 ATCCAGAGGGGTAGTTATCTGTATGGATTTTGGAGGCAAGAGTCAGATATAAATCCAT 374
301 ATCCAGAGGGGTAGTTATCTGTATGGATTTTGGAGGCAAGAGTCAGATATAAATCCAT 360
375 TCATCTCTCGCTCTCTAGCTCTGCAATCAATCAATCAATCAATCAATCAATCAATCAAT 434

Db 361 TCATCTCTCGCTCTCTAGCTCTGCAATCAATGATCCACAGGGAATTTATTACTGGAA 420
Qy 435 AGTGTCAATAGCTGTTTGGTGAGAGTCTGCGAGCTTCGGGAAGATATATTTCGATC 494
Db 421 AGTGTCAATAGCTGTTTGGTGAGAGTCTGCGAGCTTCGGGAAGATATATTTCGATC 480
Qy 495 TGTCAAGAAATATTACTCTCTCAGAACCCCGAGGAGTAGACTTCTCTAGAAATGTGCAAGAA 554
Db 481 TGTCAAGAAATATTACTCTCTCAGAACCCCGAGGAGTAGACTTCTCTAGAAATGTGCAAGAA 540
Qy 555 GCTGAGAAAGATATTAATCTCTGGGTCAAGACTCAAAACCAAGCAAAATCCCAAACTTG 614
Db 541 GCTGAGAAAGATATTAATCTCTGGGTCAAGACTCAAAACCAAGCAAAATCCCAAACTTG 600
Qy 615 TTACCTGAAGGTTCTGTAGATGGGATACAGAGATGGTCTCTGTGAATGCTGTACTTTC 674
Db 601 TTACCTGAAGGTTCTGTAGATGGGATACAGAGATGGTCTCTGTGAATGCTGTACTTTC 660
Qy 675 AAAGGAAAGTGGAAAACTCCATTTGAGAGAAACTAAATGGGCTTTATCTCTTCGGTGA 734
Db 661 AAAGGAAAGTGGAAAACTCCATTTGAGAGAAACTAAATGGGCTTTATCTCTTCGGTGA 720
Qy 735 AACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGGTGGAAGCTAAACATTCGA 794
Db 721 AACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGGTGGAAGCTAAACATTCGA 780
Qy 795 TACATAGAAGACCTAAAGGCTCAGATTTCTAGAACTCCCATATCTGTGAGATGTTAGCATG 854
Db 781 TACATAGAAGACCTAAAGGCTCAGATTTCTAGAACTCCCATATCTGTGAGATGTTAGCATG 840
Qy 855 TTCTTTGTTCTCCAGATGAAATTCGCGATGTCTCCACTGGCTGGAGCTCTCGAAAGT 914
Db 841 TTCTTTGTTCTCCAGATGAAATTCGCGATGTCTCCACTGGCTGGAGCTCTCGAAAGT 900
Qy 915 GAAATAACTATGACAACTCAACAAAGTGAACAGCAAGCAAAAGCAAAATGGCTGAAGATGAA 974
Db 901 GAAATAACTATGACAACTCAACAAAGTGAACAGCAAGCAAAAGCAAAATGGCTGAAGATGAA 960
Qy 975 GTTGAGGTATACATATATATATATGAGAGATTAATTAATCAAGATCCATTCATCTG 1034
Db 961 GTTGAGGTATACATATATATATGAGAGATTAATTAATCAAGATCCATTCATCTG 1020
Qy 1035 AAAAGCATGGGATGAGGAGGAGCGCTTCAACAGGAGCGGCGCAATTTCTCAGGATGTCG 1094
Db 1021 AGAAGCATGGGATGAGGAGGAGCGCTTCAACAGGAGCGGCGCAATTTCTCAGGATGTCG 1080
Qy 1095 GAGAGGAATGACCTGTTCTTCTGAAGTGTTCACCAAGCCATGGTGGATGTAATGAG 1154
Db 1081 GAGAGGAATGACCTGTTCTTCTGAAGTGTTCACCAAGCCATGGTGGATGTAATGAG 1140
Qy 1155 GAGGCACTGAAGCAGCGCTGSCACAGAGGTGTTATGACAGGAGAACTGGACATGGA 1214
Db 1141 GAGGCACTGAAGCAGCGCTGSCACAGAGGTGTTATGACAGGAGAACTGGACATGGA 1200
Qy 1215 GGCCCAACAGTTTGTGCGAGATCATCCGCTTTCTTTTCTTATATGATAAGATAACCAAG 1274
Db 1201 GGCCCAACAGTTTGTGCGAGATCATCTTTTCTTTTCTTATATGATAAGATAACCAAC 1260
Qy 1275 TGCATTTATTTTTCGCGAGATTTGCTCACCTTAACTTAACTGAGGTGCTGCTCTCGCAA 1334
Db 1261 TGCATTTATTTTTCGCGAGATTTTCTCACCTTAACTTAACTGAGGTGCTGCTCTCGCAA 1320
Qy 1335 AGATTTTCTAGATGAGCTGTGTGCTCAGAAATGCTATTTCATAATGGCCAAAATTTAG 1394
Db 1321 AGATTTTCTAGATGAGCTGTGTGCTCAGAAATGCTATTTCATAATGGCCAAAATTTAG 1380
Qy 1395 AGATGTTTTCTACATATTTCTGCTCTCTGAAACAACTTCTGCTACCCCACTAAATAAAAC 1454
Db 1381 AGATGTTTTCTACATATTTCTGCTCTCTGAAACAACTTCTGCTACCCCACTAAATAAAAC 1440
Qy 1455 ACAGAAATTAATAGCAATTTGCTATATTAACATGACAAACCTTATTAATCAATTTGCTCT 1514
Db 1441 ACAGAAATTAATAGCAATTTGCTATATTAACATGACAAACCTTATTAATCAATTTGCTCT 1500

```
2Y 1515 CTAAATGGGATCATGCCGATTTAGATTTTCTACTATCAGTTTATTTTATAACATTA 1574
Db 1501 CTAAATGGGATCATGCCGATTTAGATTTTCTACTATCAGTTTATTTTATAACATTA 1560
2Y 1575 ACTTTTACTTTTATTTATTTATATAATAGTGGTGAATTTTAAATTTATTCCTCACTG 1634
Db 1561 ACTTTTACTTTTATTTATTTATATAATAGTGGTGAATTTTAAATTTATTCCTCACTG 1620
2Y 1635 CCTATTTAATCTAGCTAATAAGTTTATAGCAAGCAGATGATCTGTTAATTTTCTATCTAAT 1694
Db 1621 CCTATTTAATCTAGCTAATAAGTTTATAGCAAGCAGATGATCTGTTAATTTTCTATCTAAT 1680
2Y 1695 AAATGCCCTTTAATTTCTCATATAAGAAATAAGTAGGATATCCCTCCCATGCTCTCTGT 1754
Db 1681 AAATGCCCTTTAATTTCTCATATAAGAAATAAGTAGGATATCCCTCCCATGCTCTCTGT 1740
2Y 1755 AATAAATATCTGGAATAAACAATAAGCAATATATATGTTATGTCATTTCTAGA 1814
Db 1741 AATAAATATCTGGAATAAACAATAAGCAATATATGTTATGTCATTTCTAGA 1800
2Y 1815 AATACATAACATATATATGCTGTATCTTATATTTCAATTCGAAGTATATAAT 1868
Db 1801 AATACATAACATATATATGCTGTATCTTATATTTCAATTCGAAGTATATAAT 1854

RESULT 13
LOCUS I12549 2409 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 18 from patent US 5426044.
ACCESSION I12549
VERSION I12549.1 GI:9099933
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2409)
AUTHORS Antalis,T.M., Barnes,T.M., Clark,M.A., Devine,P.L., Goss,N.H. and
Lehrbach,P.R.
TITLE Minactivin compositions and antibodies to minactivin
JOURNAL Patent: US 5426044-A 18 20-JUN-1995;
FEATURES Location/Qualifiers
source 1..2409
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 96.8%; Score 1847.6; DB 6; Length 2409;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 15 GTCAGACAGCACTCAGAGAATAACAGAGAAACAACAGAGATTGAAACAATGGAGGATCTT 74
Db 1 GTCAGACAGCACTCAGAGAATAACAGAGAAACAACAGAGATTGAAACAATGGAGGATCTT 60
QY 75 TGTGTGGGAAACACATCTTTGGCTCAATTTTCAAGCATCTGGCAAGCAAGCGCC 134
Db 61 TGTGTGGGAAACACATCTTTGGCTCAATTTTCAAGCATCTGGCAAGCAAGCGCC 120
QY 135 ACCCGAAGCTCTTCTCTCCCATGAGCATCTCGTCCACCATGGCCATGGTCTACATG 194
Db 121 ACCCGAAGCTCTTCTCTCCCATGAGCATCTCGTCCACCATGGCCATGGTCTACATG 180
QY 195 GGTCCAGGGGAGCAGCCGAGACCATGCGCAAGGTGCTTCAAGTTAATGAAGTGGGA 254
Db 181 GGTCCAGGGGAGCAGCCGAGACCATGCGCAAGGTGCTTCAAGTTAATGAAGTGGGA 240
QY 255 GCCAATGCAGTTACCCCATGATCTCCAGAACTTTTACCAGCTGTGGTTTCATGCAGCAG 314
Db 241 GCCAATGCAGTTACCCCATGATCTCCAGAACTTTTACCAGCTGTGGTTTCATGCAGCAG 300
QY 315 ATCCCAAGGAGTATCTCTGATGCGATTTCAGGACCAAGCTGCAGATAAATCCAT 1454
Db 1440 ATCCCAAGGAGTATCTCTGATGCGATTTCAGGACCAAGCTGCAGATAAATCCAT 1440
```

```
Db 301 ATCCCAAGGAGTATCTCTGATGCGATTTCAGGACCAAGCTGCAGATAAATCCAT 360
QY 375 TCATCTCTCCGCTCTCTCAGCTCTCAATCAATGCATCCACAGGGAAATTTTACTCGAA 434
Db 361 TCATCTCTCCGCTCTCTCAGCTCTCAATCAATGCATCCACAGGGAAATTTTACTCGAA 420
QY 435 AGTGTCAATAAGCTGTTTGGTGAGAGTCTGGAGCTTCGGGAGAAATATATTCGATC 494
Db 421 AGTGTCAATAAGCTGTTTGGTGAGAGTCTGGAGCTTCGGGAGAAATATATTCGATC 480
QY 495 TGTCAAGAAATATTACTCTCTCAGAACCCCGAGGAGTAGACTTCTAGAAATGTGCAAGAA 554
Db 481 TGTCAAGAAATATTACTCTCTCAGAACCCCGAGGAGTAGACTTCTAGAAATGTGCAAGAA 540
QY 555 GCTAGAAAAAGATTAATTTCTTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAACTTG 614
Db 541 GCTAGAAAAAGATTAATTTCTTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAACTTG 600
QY 615 TTACTCTGAAGGTTCTGTAGATGGGATACAGGATGTTCTTGGTGAATGCTGTACTTTC 674
Db 601 TTACTCTGAAGGTTCTGTAGATGGGATACAGGATGTTCTTGGTGAATGCTGTACTTTC 660
QY 675 AAAGGAAAGTGGAAAAACTCCATTTCGAGAAACTAAATGGGCTTTATCTTCCGTGTA 734
Db 661 AAAGGAAAGTGGAAAAACTCCATTTCGAGAAACTAAATGGGCTTTATCTTCCGTGTA 720
QY 735 AACTCGGCTCAGCGCACACCTCTACAGATGATGATCTTGGTGAAGCTAAACATTGGA 794
Db 721 AACTCGGCTCAGCGCACACCTCTACAGATGATGATCTTGGTGAAGCTAAACATTGGA 780
QY 795 TACATAGAGACCTAAAGGCTCAGATTTCTAGAACTCCCATATGCTGAGATGTTAGCATG 854
Db 781 TACATAGAGACCTAAAGGCTCAGATTTCTAGAACTCCCATATGCTGAGATGTTAGCATG 840
QY 855 TTCTTGTGCTTCCAGATGAAATTCGCGATGTTCCATCTGGTTGGAGCTCTCGAAAGT 914
Db 841 TTCTTGTGCTTCCAGATGAAATTCGCGATGTTCCATCTGGTTGGAGCTCTCGAAAGT 900
QY 915 GAAATAACCTATGACAAACTCAACAAGTGGACAGCAAGCAACAAATGGCTGAAGATGAA 974
Db 901 GAAATAACCTATGACAAACTCAACAAGTGGACAGCAAGCAACAAATGGCTGAAGATGAA 960
QY 975 GTTAGGTATACATACCCAGTTCAAAATTAGAGAGCAATTATGAACTCAGATCCATTTG 1034
Db 961 GTTAGGTATACATACCCAGTTCAAAATTAGAGAGCAATTATGAACTCAGATCCATTTG 1020
QY 1035 AAAAGCATGGCATGGAGGACGCTTCAACAAGGAGCGGCCAATTTCTCAGGGATGTCG 1094
Db 1021 AAAAGCATGGCATGGAGGAGCGCTTCAACAAGGAGCGGCCAATTTCTCAGGGATGTCG 1080
QY 1095 GAGAGAAATGACCTGTTTCTTCTGAAGTGTTCACCAAGCCATGGTGGATGTAATGAG 1154
Db 1081 GAGAGAAATGACCTGTTTCTTCTGAAGTGTTCACCAAGCCATGGTGGATGTAATGAG 1140
QY 1155 GAGGCACTGAAGCAGCCGCTGGCACAGGAGGTGTTATGACAGGAGAACTGGACATGGA 1214
Db 1141 GAGGCACTGAAGCAGCCGCTGGCACAGGAGGTGTTATGACAGGAGAACTGGACATGGA 1200
QY 1215 GGGCCACAGTTTGTGCGAGATCATCGTTTCTTCTTATTTATGCAATAAGATACCAAG 1274
Db 1201 GGGCCACAGTTTGTGCGAGATCATCGTTTCTTCTTATTTATGCAATAAGATACCAAG 1260
QY 1275 TGCATTTTATTTTTCGGCAGATTTTGTCTACCTCAAAACTAAAGCGTGTCTTTCGAAA 1334
Db 1261 TGCATTTTATTTTTCGGCAGATTTTCTCACCTCAAAACTAAAGCGTGTCTTTCGAAA 1320
QY 1335 AGATTTTCTAGATGAGCTGTGCGCTCAGAAATGCTATTTCAAATTCGCCAAAAATTTAG 1394
Db 1321 AGATTTTCTAGATGAGCTGTGCGCTCAGAAATGCTATTTCAAATTCGCCAAAAATTTAG 1380
QY 1395 AGATGTTTCTACATATTTTGTCTCTTCTGAAACAACTTCTGCTACCCCACTAAATAAAAC 1454
Db 1381 AGATGTTTCTACATATTTTGTCTCTTCTGAAACAACTTCTGCTACCCCACTAAATAAAAC 1440
```

```
Y 1455 ACAGAAATAATTAGACAATTGCTCTATTATAAATGACAAACCTATTAAATCAATTTGGTCTT 1514
b 1441 ACAGAAATAATTAGACAATTGCTCTATTATAAATGACAAACCTATTAAATCAATTTGGTCTT 1500
Y 1515 CTAAATGGGATCATGCCATTTAGATTTCCTTACTATCAGTTTATTTTATTAACATTA 1574
b 1501 CTAAATGGGATCATGCCATTTAGATTTCCTTACTATCAGTTTATTTTATTAACATTA 1560
Y 1575 ACTTTTACTTTGTTTATTATTATTATTATATATATATGTTGAGTTTATTTAAATTTATGTCACCTG 1634
b 1561 ACTTTTACTTTGTTTATTATTATTATTATATATATGTTGAGTTTATTTAAATTTATGTCACCTG 1620
Y 1635 CCTATTTAATGCTAATAAAGTTATAGAGGACATGATCTGTTAAATTTCCCTATCTAAT 1694
b 1621 CCTATTTAATGCTAATAAAGTTATAGAGGACATGATCTGTTAAATTTCCCTATCTAAT 1680
Y 1695 AATGCTTTAATTTGTTCTCTATATAGAGAAATAGTAGGTATCCCTCCATGCCCTTCTGT 1754
b 1681 AATGCTTTAATTTGTTCTCTATATAGAGAAATAGTAGGTATCCCTCCATGCCCTTCTGT 1740
Y 1755 AATAAATATCTGGAAGAAACATTAACCAATAGGCAAAATATATGTTATGTCATTTCTAGA 1814
b 1741 AATAAATATCTGGAAGAAACATTAACCAATAGGCAAAATATATGTTATGTCATTTCTAGA 1800
Y 1815 AATACATAACACATATATCTGCTATCTTATCTTATATTCATTCAGATATATAAT 1868
b 1801 AATACATAACACATATATCTGCTATCTTATCTTATATTCATTCAGATATATAAT 1854

:RESULT 14
LOCUS 10503
DEFINITION Artificial sequence for PAI-2.
ACCESSION A10503
VERSION A10503.1 GI:413574
KEYWORDS
SOURCE
ORGANISM
FEATURES
REFERENCE
AUTHORS
JOURNAL
FEATURES
source
1..1675
/organism="synthetic construct"
/mol_type="unassigned RNA"
/db_xref="taxon:32630"
1..1086
/note="Protein sequence is in conflict with the conceptual translation"
/codon_start=1
/transl_table=1
/product="PAI-2"
/protein_id="CAA00911.1"
/db_xref="GI:4526844"
/db_xref="REMITEMBL:CAA00911"
/translation="AKVLFQNEVGANAVTPMTPEPTSCGFNQOIKGSPDAILQAO
AADKHSFRSLSSAINASTNGYLLESVNKLFGEKSASFREBYIRLCQKYISEPOAV
DFLECAEARKNSWVOTQTKIPNLLPEGSVDGTRMVLNVAVYFKGKWTPEK
KLNGLYPRVNSAORTPVQMYLRKLANIGYIEDLKAQILELPYAGDVSMFLLLPDEI
ADVSTGLELLESEITVDKLNKWSKDKMAEDEVEVYIPOKLEEHYELRLSILRSMGME
DAFNKGRANFGSMERNDLFLSEVFQHMVDVNEEGTEAAAGTGGVMTGRTGHGQPOF
VADHPFLFLIMHKIINCILPFRSP"

CDS
1..1675
/organism="synthetic construct"
/mol_type="unassigned RNA"
/db_xref="taxon:32630"
1..1086
/note="Protein sequence is in conflict with the conceptual translation"
/codon_start=1
/transl_table=1
/product="PAI-2"
/protein_id="CAA00911.1"
/db_xref="GI:4526844"
/db_xref="REMITEMBL:CAA00911"
/translation="AKVLFQNEVGANAVTPMTPEPTSCGFNQOIKGSPDAILQAO
AADKHSFRSLSSAINASTNGYLLESVNKLFGEKSASFREBYIRLCQKYISEPOAV
DFLECAEARKNSWVOTQTKIPNLLPEGSVDGTRMVLNVAVYFKGKWTPEK
KLNGLYPRVNSAORTPVQMYLRKLANIGYIEDLKAQILELPYAGDVSMFLLLPDEI
ADVSTGLELLESEITVDKLNKWSKDKMAEDEVEVYIPOKLEEHYELRLSILRSMGME
DAFNKGRANFGSMERNDLFLSEVFQHMVDVNEEGTEAAAGTGGVMTGRTGHGQPOF
VADHPFLFLIMHKIINCILPFRSP"

ORIGIN
Query Match 87.3%; Score 1665.4; DB 6; Length 1675;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1669; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Y 225 GCCAAGGTGCTTCAGTTTAATGAAGTGGAGGCAATGCAAGTACCCCACTGCTCAGAG 284
b 1 GCCAAGGTGCTTCAGTTTAATGAAGTGGAGGCAATGCAAGTACCCCACTGCTCAGAG 60
```

```
QY 285 AACTTTACCAGCTCTGGGTTTCATGCAGCAGATCCAGAGGTAGTTTATCCTGATGCGATT 344
DB 61 AACTTTACCAGCTCTGGGTTTCATGCAGCAGATCCAGAGGTAGTTTATCCTGATGCGATT 120
QY 345 TTCGAGCACAAGCTCGAGATATAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATC 404
DB 121 TTGCGAGCACAAGCTCGAGATATAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATC 180
QY 405 AATGCATCCACAGGGAATTTATTACTGAAAGTGTCATTAAGCTGTTTGGTGAAGTCT 464
DB 181 AATGCATCCACAGGGAATTTATTACTGAAAGTGTCATTAAGCTGTTTGGTGAAGTCT 240
QY 465 GCGAGCTTCGCGGGAAGATATATTCGACTCTCTCAGAAATATTACTCTCAGAACCCCGAG 524
DB 241 GCGAGCTTCGCGGGAAGATATATTCGACTCTCTCAGAAATATTACTCTCAGAACCCCGAG 300
QY 525 GCAGTAGACTTCCTAGATGTGCAGAAAGCTAGAAAAGATTAATTCCTGGGTCAAG 584
DB 301 GCAGTAGACTTCCTAGATGTGCAGAAAGCTAGAAAAGATTAATTCCTGGGTCAAG 360
QY 585 ACTCAAAACCAAGGCAAAATCCCAAACCTGTTTACCTGAAAGTTCTGTAGATGGGATACC 644
DB 361 ACTCAAAACCAAGGCAAAATCCCAAACCTGTTTACCTGAAAGTTCTGTAGATGGGATACC 420
QY 645 AGGATGCTCTGCTGATGCTGCTACTTCAAGGAAGTGGAAAACCTCCATTTGAGAAG 704
DB 421 AGGATGCTCTGCTGATGCTGCTACTTCAAGGAAGTGGAAAACCTCCATTTGAGAAG 480
QY 705 AACTTAATGGGCTTTATCTCTTCGCTGTAATCTGGCTCAGCGCACACCTGTACAGATG 764
DB 481 AACTTAATGGGCTTTATCTCTTCGCTGTAATCTGGCTCAGCGCACACCTGTACAGATG 540
QY 765 ATGTACTTGCCTGAAAAGCTAAACATTGGATACATAGAAGACCTTAAAGGCTCAGATTCTA 824
DB 541 ATGTACTTGCCTGAAAAGCTAAACATTGGATACATAGAAGACCTTAAAGGCTCAGATTCTA 600
QY 825 GAATCCCATATGCTGGAGATGTAGCATGTTCTGCTTCCAGATGAATTTGCCGAT 884
DB 601 GAATCCCATATGCTGGAGATGTAGCATGTTCTGCTTCCAGATGAATTTGCCGAT 660
QY 885 GTGTCACACTGGCTTGGAGCTGCTGGAAGTGAATAACCTATGACAAACCTCAACAGTGG 944
DB 661 GTGTCACACTGGAGCTGCTGGAAGTGAATAACCTATGACAAACCTCAACAGTGG 720
QY 945 ACCAGCAAGACAAAATGGCTGAAGTGAAGTGAAGTATACATACCCAGTTCAAATTA 1004
DB 721 ACCAGCAAGACAAAATGGCTGAAGTGAAGTGAAGTATACATACCCAGTTCAAATTA 780
QY 1005 GAAGAGCATTTAAGACTCAGATCCATTCTGAAAAGCATGGGCATGGAGGACGCTTCAAC 1064
DB 781 GAAGAGCATTTAAGACTCAGATCCATTCTGAAAAGCATGGGCATGGAGGACGCTTCAAC 840
QY 1065 AAGGACCGGGCCAAATTTCTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTCTGAAGTG 1124
DB 841 AAGGACCGGGCCAAATTTCTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTCTGAAGTG 900
QY 1125 TTCACCAAGCATGCTGGATGTGAATGAGGAGGACCTGAAGCAGCGCTGGCACAGGA 1184
DB 901 TTCACCAAGCATGCTGGATGTGAATGAGGAGGACCTGAAGCAGCGCTGGCACAGGA 960
QY 1185 GGTGTATTACAGGGAGAACTGGACATGGAGGCCCAAGTTTGTGGCAGATCATCCGTTT 1244
DB 961 GGTGTATTACAGGGAGAACTGGACATGGAGGCCCAAGTTTGTGGCAGATCATCCGTTT 1020
QY 1245 CTTTTTCTTATTATGCATAAGATAACCAAGTGCATTTTATTTTTCGGCAGATTTTGTCTCA 1304
DB 1021 CTTTTTCTTATTATGCATAAGATAACCAAGTGCATTTTATTTTTCGGCAGATTTTCTCA 1080
QY 1305 CCTTAAACTAAGCGTGTCTCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAG 1364
DB 1081 CCTTAAACTAAGCGTGTCTCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAG 1140
```



```

2Y 975 GTTGGGTTATACATACCCAGTTCAAAATTAGAAAGAGCATTATGAACTCAGATCCATTCTG 1034
Db 961 GTTGGGTTATACATACCCAGTTCAAAATTAGAAAGAGCATTATGAACTCAGATCCATTCTG 1020
2Y 1035 AAAAGCATGGGCATGGAGGACCGCTTCAACAAAGGGACGGGCCAAATTTCTCAGGGATGTCG 1094
Db 1021 AGAAGCATGGGCATGGAGGACCGCTTCAACAAAGGGACGGGCCAAATTTCTCAGGGATGTCG 1080
2Y 1095 GAGAGGAATGACCTGTTCTTTCTGAGTGTTCACCAAGCCATGCGTGGATGTGAATGAG 1154
Db 1081 GAGAGGAATGACCTGTTCTTTCTGAGTGTTCACCAAGCCATGCGTGGATGTGAATGAG 1140
2Y 1155 GAGGCACTGAAGCAGCGCTGGCAGAGAGTGTATGACAGGGAGAACTGGACATGGA 1214
Db 1141 GAGGCACTGAAGCAGCGCTGGCAGAGAGTGTATGACAGGGAGAACTGGACATGGA 1200
2Y 1215 GGCCACAGTTTGTGGCAGATCATCGGTTCTTTCTTATTATGATAGATAACCAAG 1274
Db 1201 GGCCACAGTTTGTGGCAGATCATCGGTTCTTTCTTATTATGATAGATAACCAAC 1260
2Y 1275 TGCAATTTTATTTTTCGGCAGATTTTCTCACCCCTAAAACTAAGCGTGTCTCTTGCAAA 1334
Db 1261 TGCAATTTTATTTTTCGGCAGATTTTCTCACCCCTAAAACTAAGCGTGTCTCTTGCAAA 1320
2Y 1335 AGATTTTGTAGATGAGCTGTGCTCAGATTGCTATTTCAAATGCCAABAAATTTAG 1394
Db 1321 AGATTTTGTAGATGAGCTGTGCTCAGATTGCTATTTCAAATGCCAABAAATTTAG 1380
2Y 1395 AGATGTTTTCTACATATTTCTGCTCTTCTGAACAACCTCTGCTACCCACTAAATAAAAC 1454
Db 1381 AGATGTTTTCTACATATTTCTGCTCTTCTGAACAACCTCTGCTACCCACTAAATAAAAC 1440
2Y 1455 ACAGAAATAATTAGACAAATGTCTATTATAACATGACACCCCTATTAAATCATTTGGTCTT 1514
Db 1441 ACAGAAATAATTAGACAAATGTCTATTATAACATGACACCCCTATTAAATCATTTGGTCTT 1500
2Y 1515 CTAAATGGGATCATGCCCAATTTAGATTTTCTTACTATCAGTTTATTTTATAACATTA 1574
Db 1501 CTAAATGGGATCATGCCCAATTTAGATTTTCTTACTATCAGTTTATTTTATAACATTA 1560
2Y 1575 ACTTTTACTTTGTTTATTTATTTATATTAATGTTGAGTTTATAATTA 1624
Db 1561 ACTTTTACTTTGTTTATTTATTTATATAATGTTGAGTTTATAATTA 1610

```

Search completed: March 11, 2004, 03:36:22
Job time : 4868.51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

DM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 08:01:57 ; Search time 464.37 Seconds

(without alignments)
17454.974 Million cell updates/sec

Title: US-10-084-817-27

Perfect score: 1908

Sequence: 1 gagagcattgcccgtcaga.....aacaaaaataaaaaaagg 1908

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001s:*

5: Geneseqn2001bs:*

6: Geneseqn2002s:*

7: Geneseqn2003as:*

8: Geneseqn2003bs:*

9: Geneseqn2003cs:*

10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1897	99.4	1897	8	ADA24501	ADA24501	Human cdn
2	1896	99.4	1896	4	AAH33195	AAH33195	Human col
3	1896	98.8	1896	1	AAH33195	AAH33195	Sequence
4	1877.2	98.4	1910	2	AAQ12590	AAQ12590	Placental
5	1872.6	98.1	1900	9	ADD18727	ADD18727	Human dis
6	1846	96.8	2409	1	AAH70474	AAH70474	Sequence
7	1750	91.7	2424	1	AAH91177	AAH91177	Sequence
8	1667	87.4	1675	1	AAH80060	AAH80060	Sequence
9	1405.6	73.7	1412	2	AAQ10849	AAQ10849	Encodes P
10	1402.6	73.5	1512	2	AAQ11819	AAQ11819	Plasminog
11	1341	70.3	1482	2	AAQ11820	AAQ11820	Plasminog
12	1249.4	65.5	1328	2	AAQ11128	AAQ11128	PreA-plas
13	1242.8	65.1	1340	2	AAQ11129	AAQ11129	PreB-PAI-
14	1242	65.1	1290	2	AAQ21119	AAQ21119	PAI-2 in
15	1014	53.1	1024	7	ABZ83797	ABZ83797	Toxicolog
16	984.8	51.6	1199	6	ABL57801	ABL57801	Oesophagu
17	567.8	29.8	601	2	AAH88579	AAH88579	Human chr
18	560	29.4	566	4	AAH56554	AAH56554	Human cdn
19	525.2	27.5	661	3	AAH72490	AAH72490	Single nu
20	525.2	27.5	661	3	AAH72484	AAH72484	Single nu
21	525.2	27.5	661	3	AAH72487	AAH72487	Single nu
22	524	27.5	661	3	AAH72451	AAH72451	Single nu
23	523.6	27.4	661	3	AAH72454	AAH72454	Single nu

ALIGNMENTS

RESULT 1

ADA24501

ID ADA24501 standard; cdna; 1897 BP.

XX AC ADA24501;

XX DT 20-NOV-2003 (first entry)

XX DE Human cDNA differential expressed in adipose tissue, INCYTE063646CBL.

XX KW ss; differential expression; adipose tissue; cytostatic; hypotensive; antiarteriosclerotic; antidiabetic; anorectic; gene therapy;

XX KW peroxisome proliferator-activated receptor gamma; PPARGgamma;

XX KW diabetes mellitus; obesity; hypertension; atherosclerosis; breast cancer;

XX KW prostate cancer; colon cancer; polycystic ovarian syndrome.

XX OS Homo sapiens.

XX FN US2003096272-A1.

XX PD 22-MAY-2003.

XX PF 29-JUL-2002; 2002US-00208408.

XX PR 30-JUL-2001; 2001US-0308969P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Schebye XM;

XX DR WPI; 2003-606416/57.

XX PT New combination comprising several cDNAs, useful for preparing a composition for diagnosing or treating diabetes mellitus, obesity,

XX PT hypertension, atherosclerosis, or cancer of the breast, prostate or colon.

XX PS Claim 1; Page 39-40; 84pp; English.

XX CC Then invention relates to a new combination comprising 55 cDNAs (ADA24485 -ADA24539) or their complements that are differentially regulated in an adipose sample. Also included are detecting differential expression of one or more cDNAs in a sample containing nucleic acids, screening several molecules or compounds to identify a ligand that specifically binds a cDNA, a vector comprising the cDNA, a host cell containing the vector, producing a protein, screening several molecules or compounds, producing

24	523.6	27.4	561	3	AAC72496	AAC72496	Single nu
25	498.8	26.1	503	4	AAH56442	AAH56442	Human cdn
26	475.6	24.9	482	4	AAH35183	AAH35183	Probe #38
27	475.6	24.9	482	4	AAH29227	AAH29227	Human bon
28	475.6	24.9	482	4	AAH28847	AAH28847	Human liv
29	315.4	16.5	317	4	AAH48342	AAH48342	Probe #17
30	315.4	16.5	317	4	AAH42268	AAH42268	Human bon
31	315.4	16.5	317	4	AAH41876	AAH41876	Human liv
32	307.4	16.1	316	2	AAH88576	AAH88576	Human chr
33	279	14.6	1519	8	AAH57845	AAH57845	Human cdn
34	274.4	14.4	314	4	AAH56562	AAH56562	Human cdn
35	274	14.4	1278	8	AAH57843	AAH57843	Human cdn
36	272.4	14.3	1278	6	AAH54338	AAH54338	CDNA enco
37	272.4	14.3	1278	7	AAH49851	AAH49851	Human ser
38	271	14.2	1200	4	AAH01980	AAH01980	Human NOV
39	269	14.1	1185	6	AAH90345	AAH90345	Human pol
40	268.4	14.1	1316	2	AAH14255	AAH14255	Human ela
41	268.4	14.1	1316	6	AAH69984	AAH69984	Pancreas
42	268.4	14.1	1316	6	AAH84121	AAH84121	Human cdn
43	268.4	14.1	1430	7	AAH56112	AAH56112	Horse sig
44	268.4	14.1	1458	6	AAH35236	AAH35236	Human gen
45	268.4	14.1	1550	9	AAH49056	AAH49056	Human NOV

CC an antibody and the isolated antibody. The cDNAs comprise sequences which
 CC are upregulated or downregulated in response to peroxisome proliferator-
 CC activated receptor gamma (PPARGgamma) agonist. The combination comprising
 CC several cDNAs is useful for preparing a composition for diagnosis or
 CC treating diabetes mellitus, obesity, hypertension, atherosclerosis,
 CC cancer of the breast, prostate or colon, or polycystic ovarian syndrome.
 XX
 SQ Sequence 1897 BP; 595 A; 387 C; 382 G; 533 T; 0 U; 0 Other;

Query Match 99.4%; Score 1897; DB 8; Length 1897;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1897; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GAGGAGCATTTGCCGTCAGACAGCAACTCAGAGATTAACACAGAGACCAACAGATTGAAA	60
DB	1	GAGGAGCATTTGCCGTCAGACAGCAACTCAGAGATTAACACAGAGACCAACAGATTGAAA	60
QY	61	CAATGAGGATCTTTGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGG	120
DB	61	CAATGAGGATCTTTGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGG	120
QY	121	CAAAAGCAAGCCCAACAGCAACTCTCTCTCCCATGAGCATCTCGTCCACCATGG	180
DB	121	CAAAAGCAAGCCCAACAGCAACTCTCTCTCCCATGAGCATCTCGTCCACCATGG	180
QY	181	CCATGCTCTACATGGGCTCCAGGGGCGACACCGAAGACAGATGCCCAAGGTGCTTCAGT	240
DB	181	CCATGCTCTACATGGGCTCCAGGGGCGACACCGAAGACAGATGCCCAAGGTGCTTCAGT	240
QY	241	TTAATGAAGTGGGAGCCAAATGCACTTACCCCATGACTCCAGAGAACTTTACCAAGCTGTG	300
DB	241	TTAATGAAGTGGGAGCCAAATGCACTTACCCCATGACTCCAGAGAACTTTACCAAGCTGTG	300
QY	301	GGTTTCATGACAGCAGATCCAGAGGGTATGTTATCTGTATGCGATTTTGAGGCACAAAGCTG	360
DB	301	GGTTTCATGACAGCAGATCCAGAGGGTATGTTATCTGTATGCGATTTTGAGGCACAAAGCTG	360
QY	361	CAGATAAATCCATTCATCTCCGCTCTCAGCTCTGCAATCAATCAATCCACAGGGA	420
DB	361	CAGATAAATCCATTCATCTCCGCTCTCAGCTCTGCAATCAATCAATCCACAGGGA	420
QY	421	ATTATTACTGGAAAGTGTCAATAAGCTGTTTGGTGAAGTCTCGAGCTTCGCGGAAG	480
DB	421	ATTATTACTGGAAAGTGTCAATAAGCTGTTTGGTGAAGTCTCGAGCTTCGCGGAAG	480
QY	481	AATATATTCGACTGTGTCAGAAATATTAATCTCTCAGACCCAGGAGTCTTCCTAG	540
DB	481	AATATATTCGACTGTGTCAGAAATATTAATCTCTCAGACCCAGGAGTCTTCCTAG	540
QY	541	AATGTGCAGAAAGTAGAAGAAAGATTAATTTCTGGGTCAAGACTCAAAACCAAGGCA	600
DB	541	AATGTGCAGAAAGTAGAAGAAAGATTAATTTCTGGGTCAAGACTCAAAACCAAGGCA	600
QY	601	AAATCCCAACTTTTACCTGAAGTTCTGTAGATGGGATACAGGATGGTCTGGTGA	660
DB	601	AAATCCCAACTTTTACCTGAAGTTCTGTAGATGGGATACAGGATGGTCTGGTGA	660
QY	661	ATGCTGTCTACTCTCAAGGAAAGTGGAAACTCCATTTTGAGAGAAACTTAAATGGGCTTT	720
DB	661	ATGCTGTCTACTCTCAAGGAAAGTGGAAACTCCATTTTGAGAGAAACTTAAATGGGCTTT	720
QY	721	ATCCTTTCCGCTGTAACCTCGGCTCAGCGCACCTGTACAGATGATGTACTTTCGCTGAAA	780
DB	721	ATCCTTTCCGCTGTAACCTCGGCTCAGCGCACCTGTACAGATGATGTACTTTCGCTGAAA	780
QY	781	AGCTAAACATTTGGATACATAGAAGACCTTAAAGGCTCAGATTTCTAGAACTCCCATATGCTG	840
DB	781	AGCTAAACATTTGGATACATAGAAGACCTTAAAGGCTCAGATTTCTAGAACTCCCATATGCTG	840
QY	841	GAGATGTAGCATGTTCTTGTGCTTCAGATGAAATGCCGATGTGCTCACTGGCTTGG	900
DB	841	GAGATGTAGCATGTTCTTGTGCTTCAGATGAAATGCCGATGTGCTCACTGGCTTGG	900

RESULT 2
 AAH33195
 ID AAH33195 standard; cDNA; 1935 BP.
 XX

QY	901	AGCTGCTGAAAGTGAAATTAACCTATGACAAACTCAACAGTGCACAGCAAGCAACAAA	960
DB	901	AGCTGCTGAAAGTGAAATTAACCTATGACAAACTCAACAGTGCACAGCAAGCAACAAA	960
QY	961	TGGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAGAGCATTATGAAC	1020
DB	961	TGGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAGAGCATTATGAAC	1020
QY	1021	TCAGATCCATTTCTGAAAAGCATGGGATGAGAGACGCTTCAACAGGAGCGGCAATT	1080
DB	1021	TCAGATCCATTTCTGAAAAGCATGGGATGAGAGACGCTTCAACAGGAGCGGCAATT	1080
QY	1081	TCTCAGGGATGTCGAGAGAGGAATGACCTGTTTCTTCTGAAGTGTTCACCAAGCCATGG	1140
DB	1081	TCTCAGGGATGTCGAGAGAGGAATGACCTGTTTCTTCTGAAGTGTTCACCAAGCCATGG	1140
QY	1141	TGATGTGAATGAGGAGGCACTGAAGACGCGGTGACAGAGGTTTATGACAGGA	1200
DB	1141	TGATGTGAATGAGGAGGCACTGAAGACGCGGTGACAGAGGTTTATGACAGGA	1200
QY	1201	GAACTGCATGAGGAGGCCACAGTTTGTGCAGATCATCCGTTTCTTTTCTTATTATGC	1260
DB	1201	GAACTGCATGAGGAGGCCACAGTTTGTGCAGATCATCCGTTTCTTTTCTTATTATGC	1260
QY	1261	ATAAGATAACCAAGTGCATTTTATTTTTCGAGATTTTGTCTACCCCTAAACTAAGCGT	1320
DB	1261	ATAAGATAACCAAGTGCATTTTATTTTTCGAGATTTTGTCTACCCCTAAACTAAGCGT	1320
QY	1321	GCTGCTTCTGCAAAAAGATTTTGTAGATGAGCTGTGCGCTCAGAAATGCTATTCAAT	1380
DB	1321	GCTGCTTCTGCAAAAAGATTTTGTAGATGAGCTGTGCGCTCAGAAATGCTATTCAAT	1380
QY	1381	TGCAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTCTGACAACTCTGCTACC	1440
DB	1381	TGCAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTCTGACAACTCTGCTACC	1440
QY	1441	CACATAAATAAACAACAGAAATTAATAGACAAATGCTATTATTAACATGACAAACCTATT	1500
DB	1441	CACATAAATAAACAACAGAAATTAATAGACAAATGCTATTATTAACATGACAAACCTATT	1500
QY	1501	AATCATTTGGTCTCTTAAATGGGATCATGCCCAATTTAGATTTTCTTACTATCAGTTTA	1560
DB	1501	AATCATTTGGTCTCTTAAATGGGATCATGCCCAATTTAGATTTTCTTACTATCAGTTTA	1560
QY	1561	TTTATTATAACATTAACCTTTTACITTTGTTTATTATTATTTATATAATGGTGAGTTTAA	1620
DB	1561	TTTATTATAACATTAACCTTTTACITTTGTTTATTATTATTTATATAATGGTGAGTTTAA	1620
QY	1621	ATTATTGCTCAGTCCCTATTATTAATGAGCTAATAAGTTATAGAGCAGATGATCTGTTA	1680
DB	1621	ATTATTGCTCAGTCCCTATTATTAATGAGCTAATAAGTTATAGAGCAGATGATCTGTTA	1680
QY	1681	ATTTCTCTATCTAATAAATGCTTTTAAATTTGTTCTCATATGAAGATAAGTAGTATCCCT	1740
DB	1681	ATTTCTCTATCTAATAAATGCTTTTAAATTTGTTCTCATATGAAGATAAGTAGTATCCCT	1740
QY	1741	CCATGCCCTTCTGTAATAAATATCTGAAAACCAATTAAACCAATAGGCAATATATGTTA	1800
DB	1741	CCATGCCCTTCTGTAATAAATATCTGAAAACCAATTAAACCAATAGGCAATATATGTTA	1800
QY	1801	TGTGCACTTTCTAGAAATACATATATATGCTCTGTATCTTATATTCATTTGCAAG	1860
DB	1801	TGTGCACTTTCTAGAAATACATATATATGCTCTGTATCTTATATTCATTTGCAAG	1860
QY	1861	TATATAATAAATAACCTGCTTCCAAACCAACAAAAA	1897
DB	1861	TATATAATAAATAACCTGCTTCCAAACCAACAAAAA	1897

AC	AAH33195;	QY	243	AATGAAGTGGAGCCCAATGCAGTTACCCCATGACCTCCAGAGAACTTTTACCAGCTGTGGG	302
XX		Db	255	AATGAAGTGGAGCCCAATGCAGTTACCCCATGACCTCCAGAGAACTTTTACCAGCTGTGGG	314
DE	03-SEP-2001 (first entry)	QY	303	TTTCATGACGACATCCAGAAAGGTAGTTATCTCGATGCGATTTTTCAGGCAAGACTGCA	362
XX		Db	315	TTTCATGACGACATCCAGAAAGGTAGTTATCTCGATGCGATTTTTCAGGCAAGACTGCA	374
KW	Human; colon cancer; colon cancer antigen; diagnosis; detection;	QY	363	GATAAAATCCATTTCATCTCTCGCTCTCTCAGCTCTGCAATCAATCAATCCACAGGGAAT	422
KW	colorectal carcinoma; ss.	Db	375	GATAAAATCCATTTCATCTCTCGCTCTCTCAGCTCTGCAATCAATCCACAGGGAAT	434
OS	Homo sapiens.	QY	423	TATTACTGGAAGGTGTCATTAAGCTGTTTGGTGAGAGTCTGCGAGCTTCCGGGAAGAA	482
XX		Db	435	TATTACTGGAAGGTGTCATTAAGCTGTTTGGTGAGAGTCTGCGAGCTTCCGGGAAGAA	494
EN	WO200122920-A2.	QY	483	TATATTGCACTCTGTCAGAAATATTACTCTCAGAACCCAGGAGTAGACTTCTCTAGAA	542
XX	05-APR-2001.	Db	495	TATATTGCACTCTGTCAGAAATATTACTCTCAGAACCCAGGAGTAGACTTCTCTAGAA	554
PR	28-SEP-2000; 2000WO-US026524.	QY	543	TGTGAGAAAGCTTAGAAAAAGATTATTCCTGGTCAAGACTCAAAACCAAGGCAAA	602
PR	29-SEP-1999; 99US-0157137P.	Db	555	TGTGAGAAAGCTTAGAAAAAGATTATTCCTGGTCAAGACTCAAAACCAAGGCAAA	614
PR	03-NOV-1999; 99US-0163280P.	QY	603	ATCCCAAACTTGTACTCAAGGTTCTGTAGATGGGATACCAAGATGGTCTCTGTGTAAT	662
XX		Db	615	ATCCCAAACTTGTACTCAAGGTTCTGTAGATGGGATACCAAGATGGTCTCTGTGTAAT	674
XX	(HUMA-) HUMAN GENOME SCI INC.	QY	663	GCTGCTACTTCAAAGGAAGTGAAGAACTCCATTTGAGAAGAACTTAAATGGGCTTTAT	722
XX	Ruben SM, Barash SC, Birse CE, Rosen CA;	Db	675	GCTGCTACTTCAAAGGAAGTGAAGAACTCCATTTGAGAAGAACTTAAATGGGCTTTAT	734
XX	WPI; 2001-235357/24.	QY	723	CTTTTCCGTTGTAACCTCGCTCAGCGCACACCTGTACAGATGATGTACTTCCGTGAAAG	782
XX	P-PSDB; AAG73764.	Db	735	CTTTTCCGTTGTAACCTCGCTCAGCGCACACCTGTACAGATGATGTACTTCCGTGAAAG	794
XX	Nucleic acids encoding 4277 human colon cancer-associated polypeptides,	QY	783	CTAAACATTTGGATACATAGAGACCTTAAGGCTCAGATTTAGAACTCCCATATCTGGA	842
XX	useful for preventing, diagnosing and/or treating colorectal cancers.	Db	795	CTAAACATTTGGATACATAGAGACCTTAAGGCTCAGATTTAGAACTCCCATATCTGGA	854
XX	Claim 1; Page 2385-2386; 9803pp; English.	QY	843	GATGTTAGCATGTTCTTGTGTTCTCCAGATGAAATTCCTGATGTGTCCACTGGCTTGGAG	902
XX	AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon	Db	855	GATGTTAGCATGTTCTTGTGTTCTCCAGATGAAATTCCTGATGTGTCCACTGGCTTGGAG	914
CC	cancer-associated nucleic acid molecules (N) and proteins (P), where the	QY	903	CTGCTGGAAGTGAATTAACCTTATGACAACTCAAGTGGACCAAGCAAGCAAAATG	962
CC	proteins are collectively known as colon cancer antigens. The colon	Db	915	CTGCTGGAAGTGAATTAACCTTATGACAACTCAAGTGGACCAAGCAAGCAAAATG	974
CC	cancer antigens have cytostatic activity and can be used in gene therapy	QY	963	GCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAAGACATTTATGAATC	1022
CC	and vaccine production. N and P may be used in the prevention, diagnosis	Db	975	GCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAAGACATTTATGAATC	1034
CC	and treatment of diseases associated with inappropriate P expression. For	QY	1023	AGATCCATTTCTGAAGATGAGGAGCGCTTCAACAGGAGCGGCCAATTC	1082
CC	example, N and P may be used to treat disorders associated with decreased	Db	1035	AGATCCATTTCTGAAGATGAGGAGCGCTTCAACAGGAGCGGCCAATTC	1094
CC	expression by rectifying mutations or deletions in a patient's genome	QY	1083	TCAGGATGTCGGAGAGGAATGACCTGTTTCTTTCTGAAGTGTTCACCAAGCCATGGTG	1142
CC	that affect the activity of P by expressing inactive proteins or to	Db	1095	TCAGGATGTCGGAGAGGAATGACCTGTTTCTTTCTGAAGTGTTCACCAAGCCATGGTG	1154
CC	supplement the patient's own production of P. Additionally, N may be used	QY	1143	GATGATGATGAGAGGAGGACATGAGCAGCGCTGGGACAGAGGTGTTATCAGAGGAGA	1202
CC	to produce the colon cancer-associated Ps, by inserting the nucleic acids	Db	1155	GATGATGATGAGAGGAGGACATGAGCAGCGCTGGGACAGAGGTGTTATCAGAGGAGA	1214
CC	into a host cell and culturing the cell to express the proteins. N and P	QY	1203	ACTGGACATGAGAGGCGCCACAGATTTGTGGCAGATCATCCGTTTCTTTTATTATGAT	1262
CC	can be used in the prevention, diagnosis and treatment of colorectal	Db	1215	ACTGGACATGAGAGGCGCCACAGATTTGTGGCAGATCATCCCTTTCTTTTATTATGAT	1274
CC	carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent	QY	1263	AGATAAACCAAGTGCATTTTATTTTTCGGCAGATTTTGTCTCACCCCTAAACCTAGCGTC	1322
CC	sequences used in the exemplification of the present invention. N.B.	Db	1275	AGATAAACCAAGTGCATTTTATTTTTCGGCAGATTTTCTCCCTCAACCTAAACCTAGCGTC	1334
CC	Pages 666 to 682 and page 7053 of the sequence listing were missing at	QY	1323	TGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAGAAATGCTATTTCAAATTTG	1382
CC	time of publication, meaning no sequences are present for SEQ ID NO:1027	Db			
CC	to 1052, 7921 and 7922	QY			
XX	Sequence 1935 BP; 611 A; 393 C; 392 G; 539 T; 0 U; 0 Other;	QY			
XX	Query Match 99.4%; Score 1896; DB 4; Length 1935;	QY			
XX	Best Local Similarity 99.7%; Pred. No. 0;	QY			
XX	Matches 1899; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	QY			
QY	3 GGAGCATTCCTGTCAGACGACCACTCAGAGATTAACCAAGACACCAAGATTTGAACA	62			
Db	15 GGAGCATTCCTGTCAGACGACCACTCAGAGATTAACCAAGACACCAAGATTTGAACA	74			
QY	63 ATGGAGATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGCA	122			
Db	75 ATGGAGATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGCA	134			
QY	123 AAAGCAAGCCCAACCCAGAACCTTCTCTCCCTGGAGCATCTCGTCCACCATGGCC	182			
Db	135 AAAGCAAGCCCAACCCAGAACCTTCTCTCTCCCTGGAGCATCTCGTCCACCATGGCC	194			
QY	183 ATGGTCTACATGGGCTCCAGGGGACAGCCGAGACAGATGGCCAAAGGTGCTTCAGTTT	242			
Db	195 ATGGTCTACATGGGCTCCAGGGGACAGCCGAGACAGATGGCCAAAGGTGCTTCAGTTT	254			

311 GGTTCATGCAGCAGATCAGAGGGTAGTTATCCTGATCGGATTTTGCAGGCACAGCTG 370
361 CAGATAAAATCCATTCCTTCCTCCCTCTCAGCTCTGCAATCAATGATCCACAGGGA 420
371 CAGATAAAATCCATTCCTTCCTCCCTCTCAGCTCTGCAATCAATGATCCACAGGGA 430
421 ATATTTTACTGGAAGTGTCAATAAGCTGTTTGGTGAGAAGTCTGCGAGCTTCCGGGAAG 480
431 ATATTTTACTGGAAGTGTCAATAAGCTGTTTGGTGAGAAGTCTGCGAGCTTCCGGGAAG 490
481 AATATATTCGACTCTGTCAGAAATATTAATCTCTCAGAACCCAGGCGAGTAGACTTCCTAG 540
491 AATATATTCGACTCTGTCAGAAATATTAATCTCTCAGAACCCAGGCGAGTAGACTTCCTAG 550
541 AATCTGCAAGAAGAGCTAGAAAAAGATTAATTCCTGGGTCAAGATCAAAACCAAGGGA 600
551 AATCTGCAAGAAGAGCTAGAAAAAGATTAATTCCTGGGTCAAGATCAAAACCAAGGGA 610
601 AATATCCCAAACTTGTACTGAAGTTCGTAGATGGGATACAGAGTGGTCTGGTGA 660
611 AATATCCCAAACTTGTACTGAAGTTCGTAGATGGGATACAGAGTGGTCTGGTGA 670
661 ATGCTGTCTACTTCAAGAAAGTGAAGAACTCCATTTGAGAGAAACTAAATGGGCTTT 720
671 ATGCTGTCTACTTCAAGAAAGTGAAGAACTCCATTTGAGAGAAACTAAATGGGCTTT 730
721 ATCTTTCCGTGTAACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTCGCTGAAA 780
731 ATCTTTCCGTGTAACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTCGCTGAAA 790
781 AGCTAAACATTGGATACATAGAGACCTAAAGCTCAGATTCAGAACTCCCAATGCTG 840
791 AGCTAAACATTGGATACATAGAGACCTAAAGCTCAGATTCAGAACTCCCAATGCTG 850
841 GAGATGTTAGCATGTTCTTGTTCCTCCAGATGAAATTCGCGATGTTCCACTGGCTGG 900
851 GAGATGTTAGCATGTTCTTGTTCCTCCAGATGAAATTCGCGATGTTCCACTGGCTGG 910
901 AGCTGCTGGAAGTCAATAAATCTATGACAACTCAACAGTGCAGCAGCAAGACAAAA 960
911 AGCTGCTGGAAGTCAATAAATCTATGACAACTCAACAGTGCAGCAGCAAGACAAAA 970
961 TGGCTGAAGATGAAGTGAAGTATACATACCCAGTTCAAAATTAGAGAGCATTATGAAC 1020
971 TGGCTGAAGATGAAGTGAAGTATACATACCCAGTTCAAAATTAGAGAGCATTATGAAC 1030
1021 TCAGATCCATTCTGAAAAGCATGGGCATGGAGACGCTTCAACAGGGACGGGCCAATT 1080
1031 TCAGATCCATTCTGAGAAGCATGGGCATGGAGACGCTTCAACAGGGACGGGCCAATT 1090
1081 TCTCAGGATGCGGAGAGGAATGACTGTTCTTCTGAAAGTGTCCACCAAGCCATGG 1140
1091 TCTCAGGATGCGGAGAGGAATGACTGTTCTTCTGAAAGTGTCCACCAAGCCATGG 1150
1141 TGGATGTGAATGAGAGGGCCTGAAAGCAGCCGCTGGCAGAGAGTGTATGACAGGGA 1200
1151 TGGATGTGAATGAGAGGGCCTGAAAGCAGCCGCTGGCAGAGAGTGTATGACAGGGA 1210
1201 GAACTGGACATGGAGGCCACAGTTGTGGCAGATCATCGTTCTTCTTATTTATATGC 1260
1211 GAACTGGACATGGAGGCCACAGTTGTGGCAGATCATCGTTCTTCTTATTTATATGC 1270
1261 ATAAGATACCAAGTGCATTTTATTTTTCGGCAGATTTTCTCACCCTAAACCTAAGCGT 1320
1271 ATAAGATACCAAGTGCATTTTATTTTTCGGCAGATTTTCTCACCCTAAACCTAAGCGT 1330
1321 GCTGCTCTGCAAAAGATTTTCTAGATGAGTGTGCTCAGAAATGCTATTTCAAAT 1380
1331 GCTGCTCTGCAAAAGATTTTCTAGATGAGTGTGCTCAGAAATGCTATTTCAAAT 1390
1381 TGCCAAAAATTTAGAGATGTTTCTTACATATTTCTGCTTCTGGAACAACTTCGCTACC 1440
1391 TGCCAAAAATTTAGAGATGTTTCTTACATATTTCTGCTTCTGGAACAACTTCGCTACC 1450

QY 1441 CACTAATAAACAACAGAAATATTAGACAATTCCTATTATAACATGACAAACCCCTATT 1500
Db 1451 CACTAATAAACAACAGAAATATTAGACAATTCCTATTATAACATGACAAACCCCTATT 1510
QY 1501 AATCATTTGGTCTTCTAAAATGGGATCATGCCATTTAGATTTTCTTACTACTAGTTTA 1560
Db 1511 AATCATTTGGTCTTCTAAAATGGGATCATGCCATTTAGATTTTCTTACTACTAGTTTA 1570
QY 1561 TTTTATAACAATTAACATTTTACTTTTGTATTATTATTATTATTATATATGTTGAGTTTTTAA 1620
Db 1571 TTTTATAACAATTAACATTTTACTTTTGTATTATTATTATTATTATATATGTTGAGTTTTTAA 1630
QY 1621 ATTATTGCTCACTGCTTATTTAAATGCTAGCTATATAAGTTTATAGAAGCAGATGATCTGTTA 1680
Db 1631 ATTATTGCTCACTGCTTATTTAAATGCTAGCTATATAAGTTTATAGAAGCAGATGATCTGTTA 1690
QY 1681 ATTTCCCTATCTAATAAATGCCCTTTAAATGCTTCTAATAATGAAGAAATAAGTAGTATCCCT 1740
Db 1691 ATTTCCCTATCTAATAAATGCCCTTTAAATGCTTCTAATAATGAAGAAATAAGTAGTATCCCT 1749
QY 1741 CCATGCCCTCTGTAATAAATATATCTGGAAGAAACATTAAGCAATAGGCAATATATGTTA 1800
Db 1750 CCATGCCCTCTGTAATAAATATATCTGGAAGAAACATTAAGCAATAGGCAATATATGTTA 1809
QY 1801 TGTGCATTTCTAGAAATACATAACACATATATATGCTGTATCTTATTTCAATTTGCAAG 1860
Db 1810 TGTGCATTTCTAGAAATACATAACACATATATATGCTGTATCTTATTTCAATTTGCAAG 1869
QY 1861 TATATAATAAATAAACCTGCTTCCAAACAAACAAAAATAAAAAAA 1906
Db 1870 TATATAATAAATAAACCTGCTTCCAAACAAACAAAAATAAAAAAA 1915

RESULT 4

AAQ12590
ID AAQ12590 standard; cDNA; 1910 BP.
XX
AC AAQ12590;
XX AC
DT 25-MAR-2003 (revised)
DT 26-SEP-1991 (first entry)
XX
DE Placental PAI-2.
XX
KW Plasminogen activator inhibitor; anticoagulant; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT 56..1303
FT /*tag= a
FT polyA_signal 1439
FT /*tag= b
FT polyA_signal 1644
FT /*tag= c
FT polyA_signal 1685
FT /*tag= d
FT polyA_signal 1748
FT /*tag= e
FT polyA_signal 1859
FT /*tag= f
FT polyA_signal 1863
FT /*tag= g
FT polyA_site 1885..1910
FT /*tag= h
XX
US028534-A.
PN
XX
XX 02-JUL-1991.
XX
XX 15-SEP-1987; 87US-00097482.
XX

PR 15-SEP-1987; 87US-00097482.
 CX (UNIW) UNIV WASHINGTON.
 PA (MONS) MONSANTO CO.
 KX
 PI Sadler JE, Wun TC;
 DR WPI; 1991-215146/29.
 XX P-FSDB; AAR13007.
 XX
 XX Deoxyribonucleic acid - encoding recombinant human plasminogen activator
 PT inhibitor, used for preventing conversion of plasminogen to plasmin.
 XX
 XX Claim 1; Fig 4; 23pp; English.
 XX
 XX The sequence was obtd. from the combined sequences of lambda PAI-75.1 and
 CC lambda PAI75.15, both isolated from a placental cDNA library. A
 CC recombinant vector contg. the DNA can be used to express PAI-2 which
 CC inhibits the conversion of plasminogen to plasmin and is used as an
 CC anticoagulant. (Updated on 25-MAR-2003 to correct PA field.)
 XX
 XX Sequence 1910 BP; 611 A; 382 C; 384 G; 533 T; 0 U; 0 Other;
 SQ
 Query Match 98.4%; Score 1877.2; DB 2; Length 1910;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1885; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 9 TTGCGCGTCAGACAGCACTCAGAGATATACAGAGAACACAGATGAAACAAATGGAG 68
 DB 2 TTACCCGTCAGACAGCACTCAGAGATATACAGAGAACACAGATGAAACAAATGGAG 61
 QY 69 GATCTTTGTGTGGCAACACACACTCTTTGCCCTCAATTTATTCAAGCATCTGGCAAAAGCA 128
 DB 62 GATCTTTGTGTGGCAACACACACTCTTTGCCCTCAATTTATTCAAGCATCTGGCAAAAGCA 121
 QY 129 AGCCCAACAGCAACCTCTTCTCCCGATGAGCATCTGTCACCATGGCCATGGTC 188
 DB 122 AGCCCAACAGCAACCTCTTCTCTCCCGATGAGCATCTGTCACCATGGCCATGGTC 181
 QY 189 TACATGGGTCAGAGGAGCAGCAGAGACAGATGCGCAAGGTCTTCAGTTTAATGAA 248
 DB 182 TACATGGGTCAGAGGAGCAGCAGAGACAGATGCGCAAGGTCTTCAGTTTAATGAA 241
 QY 249 GTGGGAGCCAAATCAGTTACCCCATGACCTCCAGAGCACTTTACAGCTGTGGGTTTCATG 308
 DB 242 GTGGGAGCCAAATCAGTTACCCCATGACCTCCAGAGCACTTTACAGCTGTGGGTTTCATG 301
 QY 309 CAGCAGATCCAGAGGAGTAGTTATCTCGATGCGATTTTGAGGACAGCTGCAGATAAA 368
 DB 302 CAGCAGATCCAGAGGAGTAGTTATCTCGATGCGATTTTGAGGACAGCTGCAGATAAA 361
 QY 369 ATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCAGAGGAAATTTTA 428
 DB 362 ATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCAGAGGAAATTTTA 421
 QY 429 CTGGAAGTCTCAATAGCTTTTGGTGAAGTCTGGAGCTTCGGAGCTTCGGGAAGAAATATT 488
 DB 422 CTGGAAGTCTCAATAGCTTTTGGTGAAGTCTGGAGCTTCGGAGCTTCGGGAAGAAATATT 481
 QY 489 CGACTCTGTGAGAAATATTACTCTCGAAGCCCGAGGAGTAGACTTCTAGAAATGTGCA 548
 DB 482 CGACTCTGTGAGAAATATTACTCTCGAAGCCCGAGGAGTAGACTTCTAGAAATGTGCA 541
 QY 549 GAAGAGCTGAGAAAGAAATTAATTCCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCA 608
 DB 542 GAAGAGCTGAGAAAGAAATTAATTCCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCA 601
 QY 609 AACTGTGTAACCTGAAGGTTCTGTAGATGGGATACAGAGTGGTCTGGTGAATGCTGTC 668
 DB 602 AACTGTGTAACCTGAAGGTTCTGTAGATGGGATACAGAGTGGTCTGGTGAATGCTGTC 661
 QY 669 TACTTCAAGAGGAGTGGAAAGCACTCCATTTGAGAGAAAGCACTTAATGGGCTTTATCCTTTC 728

DB 562 TACTTCAAGAGGAGTGGAAAGCACTCCATTTGAGAGAAAGCACTAAATGGGCTTTATCCTTTC 721
 QY 729 CGTGTAAATCTGGCTCAGCGCACACTGTACAGATGATCTTCTGCTGAAGAGCTAAAC 788
 DB 722 CGTGTAAATCTGGCTCAGCGCACACTGTACAGATGATCTTCTGCTGAAGAGCTAAAC 781
 QY 789 ATTGGATACATAGAGACCTAAAGGCTCAGATTTCTAGAACTCCCATATGCTGGAGATGTT 848
 DB 782 ATTGGATACATAGAGACCTAAAGGCTCAGATTTCTAGAACTCCCATATGCTGGAGATGTT 841
 QY 849 AGCATGTTCTTTGTTCTCCAGATGAAATTTGCGATGTTGCTCCACTGGCTTGGAGCTGCTG 908
 DB 842 AGCATGTTCTTTGTTCTCCAGATGAAATTTGCGATGTTGCTCCACTGGCTTGGAGCTGCTG 901
 QY 909 GAAAGTGAATAAATCTATGACAAATCTCAACAAGTGCACGCAAGCAAAATGGCTGAA 968
 DB 902 GAAAGTGAATAAATCTATGACAAATCTCAACAAGTGCACGCAAGCAAAATGGCTGAA 961
 QY 969 GATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAGAGCATTTATGAATCAGATCC 1028
 DB 962 GATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAGAGCATTTATGAATCAGATCC 1021
 QY 1029 ATTCTGAAAGCATGGGATGAGGACGCTTCAACAAGGAGCGGCAATTTCTCAGG 1088
 DB 1022 ATTCTGAAAGCATGGGATGAGGACGCTTCAACAAGGAGCGGCAATTTCTCAGG 1081
 QY 1089 ATGTCCGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATGCTGGATGTG 1148
 DB 1082 ATGTCCGAGAGGAATGAGCTGTTTCTTCTGAAAGTGTTCACCAAGCCATGCTGGATGTG 1141
 QY 1149 AATGAGGAGGACCTGAAAGCAGCGCTGACAGAGGAGTGTATGACAGGAGAGAACTGGA 1208
 DB 1142 AATGAGGAGGACCTGAAAGCAGCGCTGACAGAGGAGTGTATGACAGGAGAGAACTGGA 1201
 QY 1209 CATGAGGAGGACCTGAGTGTGAGATCATTCGCTTTCTTTTCTTTTCTTTATATGATGATGATA 1268
 DB 1202 CATGAGGAGGACCTGAGTGTGAGATCATTCGCTTTCTTTTCTTTTCTTTATGATGATGATA 1261
 QY 1269 ACCAAGTGCATTTATTTTTCGSCAGATTTGCTCACCCCTAAAGCTAAAGCTGCTGCTTC 1328
 DB 1262 ACCAAGTGCATTTATTTTTCGSCAGATTTGCTCACCCCTAAAGCTAAAGCTGCTGCTTC 1321
 QY 1329 TCGAAAGATTTTGTAGATGAGCTGTGCTCAGATTTGCTTATTTCAATTTGCAAAAT 1388
 DB 1322 TCGAAAGATTTTGTAGATGAGCTGTGCTCAGATTTGCTTATTTCAATTTGCAAAAT 1381
 QY 1389 ATTGAGAGATGTTTCTTACATATTTCTGCTCTTCTGAAACAACTTCTGCTACCCATAAT 1448
 DB 1382 ATTGAGAGATGTTTCTTACATATTTCTGCTCTTCTGAAACAACTTCTGCTACCCATAAT 1441
 QY 1449 AAAACACAGAAATTAATGAGAAATTTAGCAATTTCTATTATACATGACAAACCTATTATCAATT 1508
 DB 1442 AAAACACAGAAATTAATGAGAAATTTAGCAATTTCTATTATACATGACAAACCTATTATCAATT 1501
 QY 1509 GGTCTTTTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATTTTATA 1568
 DB 1502 GGTCTTTTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATTTTATA 1561
 QY 1569 ACATTAACCTTTTACTTTTATTTTATTTTATTTATTAATGCTGAGTTTAAATTTATGTC 1628
 DB 1562 ACATTAACCTTTTACTTTTATTTTATTTTATTTATTAATGCTGAGTTTAAATTTATGTC 1621
 QY 1629 TCACTGCTATTTAAATGAGTAAATTAAGATTAAGAGCAGATGATCTGTTAAATTTCTTA 1688
 DB 1622 TCACTGCTATTTAAATGAGTAAATTAAGATTAAGAGCAGATGATCTGTTAAATTTCTTA 1681
 QY 1689 TCTAAATAAATGCTTTTAAATTTTCTCATAATGAAGAATAAGTAGGTATCCCTCCATGCC 1748
 DB 1682 TCTAAATAAATGCTTTTAAATTTTCTCATAATGAAGAATAAGTAGGTATCCCTCCATGCC 1741
 QY 1749 TCTGTAATTAATCTGGAAGAAACCAATTAACCAATAGGAAATATATGTTATGTCATT 1808
 DB 1742 TCTGTAATTAATCTGGAAGAAACCAATTAACCAATAGGAAATATATGTTATGTCATT 1801

1809 TCTAGAAATACATACACATATATATCTCTGTATCTTATATTTCAATTCGAACTATATAT 1868
 1802 TCTAGAAATACATACACATATATATCTCTGTATCTTATATTTCAATTCGAACTATATAT 1861
 1869 AAATAAACCTGCTTCCAAACACCAAAAAATAAAAAA 1906
 1862 AAATAAACCTGCTTCCAAACCAAGAAAAATAAAAAA 1899

RESULT 5
 DD18727
 D ADD18727 standard; DNA; 1900 BP.
 X ADD18727;
 C ADD18727;
 T 15-JAN-2004 (first entry)
 X Human disease related protein DNA sequence SeqID158.
 X human; disease state; cytostatic; antiinflammatory; ophthalmological;
 X antiarteriosclerotic; vulnery; gene therapy; angiogenesis; apoptosis;
 X hypoxia-regulated condition; tumorigenesis; glycolysis; gluconeogenesis;
 X inflammation; erythropoiesis; catecholamine synthesis; iron transport;
 X glucose transportation; catecholamine synthesis; iron transport;
 X nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
 X retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
 X inflammatory condition; wound healing; gene; ds.
 X Homo sapiens.
 X WO2003018621-A2.
 X 06-MAR-2003.
 X 23-AUG-2002; 2002WO-GB003892.
 X 23-AUG-2001; 2001GB-00020558.
 X 05-OCT-2001; 2001GB-00024037.
 X (OXFO-) OXFORD BIOMEDICA UK LTD.
 X Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
 WPI; 2003-290046/28.
 X P-PSDB; ADD18726.
 X New substantially purified polypeptide, useful for diagnosing or treating
 X a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
 X injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
 X wound healing.
 X Claim 27; SEQ ID NO 158; 424pp; English.
 X This invention relates to novel human genes and gene product which are
 X implicated in certain disease states. Compounds which modulate the
 X proteins of the invention may have cytostatic, antiinflammatory,
 X ophthalmological, antiarteriosclerotic or vulnery activities. The
 X sequences of the invention may be useful for gene therapy. The invention
 X may be useful for diagnosing or treating a hypoxia-regulated condition,
 X such as tumorigenesis, angiogenesis, apoptosis, inflammation,
 X erythropoiesis, or the biological response to hypoxia conditions
 X including processes such as glycolysis, gluconeogenesis, glucose
 X transportation, catecholamine synthesis, iron transport or nitric oxide
 X synthesis. The disease includes cancer, ischaemic conditions, reperfusion
 X injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
 X inflammatory conditions or wound healing. The present sequence is that of
 X a disease related protein encoding DNA sequence of the invention.
 X Sequence 1900 BP; 592 A; 393 C; 380 G; 535 T; 0 U; 0 Other;
 X Query Match 98.1%; Score 1872.6; DB 9; Length 1900;
 X Best Local Similarity 99.7%; Pred. No. 0;

		Matches 1886;	Conservative	0;	Mismatches	4;	Indels	1;	Gaps	1;
Qy	1	GAGGAGCATTCCCGTCCAGACAGCAGCACTCAGAGAAATACCCAGAGAACCAACAGATTGAAA	60							
Db	11	GAGGAGCATTCCCGTCCAGACAGCAGCACTCAGAGAAATACCCAGAGAACCAACAGATTGAAA	70							
Qy	61	CAATGGAGGATCTTTGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGG	120							
Db	71	CAATGGAGGATCTTTGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGG	130							
Qy	121	CAAAAGCAAGCCCAACCCAGAACCTCTTCTCTCTCCCATGGAGCATCTCTCCACCATGG	180							
Db	131	CAAAAGCAAGCCCAACCCAGAACCTCTTCTCTCTCCCATGGAGCATCTCTCCACCATGG	190							
Qy	181	CCATGGTCTACATGGGCTCCAGGGCAGCACCGAAGACCAGATGGCCAGGTCCTTCAGT	240							
Db	191	CCATGGTCTACATGGGCTCCAGGGCAGCACCGAAGACCAGATGGCCAGGTCCTTCAGT	250							
Qy	241	TTAATGAAGTGGAGCCCAATGCAGTTACCCCATGACTCCAGAGAACTTTTACCAGCTGTG	300							
Db	251	TTAATGAAGTGGAGCCCAATGCAGTTACCCCATGACTCCAGAGAACTTTTACCAGCTGTG	310							
Qy	301	GGTTCATGCAGCAGATCCAGAGGGGTAGTTATCTGATGCGATTTTGCAGGCACAAGCTG	360							
Db	311	GGTTCATGCAGCAGATCCAGAGGGGTAGTTATCTGATGCGATTTTGCAGGCACAAGCTG	370							
Qy	361	CAGATAAATCAATTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGA	420							
Db	371	CAGATAAATCAATTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGA	430							
Qy	421	ATTATTTACTGGAAGTGTCAATAAGCTCTTTGGTGAGAGTCTGGAGGCTTCCGGGAAG	480							
Db	431	ATTATTTACTGGAAGTGTCAATAAGCTCTTTGGTGAGAGTCTGGAGGCTTCCGGGAAG	490							
Qy	481	AATATTTGACTCTGTGAGAAATATTCTCTCAGAACCCAGGCGATGAGCTTCCTAG	540							
Db	491	AATATTTGACTCTGTGAGAAATATTCTCTCAGAACCCAGGCGATGAGCTTCCTAG	550							
Qy	541	AATGTGCAAGAGAGCTAGAAAAAGATTAATTTCTGGGTCAAGACTCAAAACCAAGGCA	600							
Db	551	AATGTGCAAGAGAGCTAGAAAAAGATTAATTTCTGGGTCAAGACTCAAAACCAAGGCA	610							
Qy	601	AAATCCCAACTTGTCTAGTGAAGTCTGTAGATGGGATACAGGATGCTCTGGTGA	660							
Db	611	AAATCCCAACTTGTCTAGTGAAGTCTGTAGATGGGATACAGGATGCTCTGGTGA	670							
Qy	661	ATGCTGTCTACTTCAAGGAAAGTGGAAAACTCCATTGAGAGAAACTAAATGGGCTTT	720							
Db	671	ATGCTGTCTACTTCAAGGAAAGTGGAAAACTCCATTGAGAGAAACTAAATGGGCTTT	730							
Qy	721	ATCCTTTCCGCTAAACTCGGCTCAGCGCACACCTGTACAGATGATGATCTTGGGTGAAA	780							
Db	731	ATCCTTTCCGCTAAACTCGGCTCAGCGCACACCTGTACAGATGATGATCTTGGGTGAAA	790							
Qy	781	AGCTAAACATTGGATACATAGAACCTTAAGGCTCAGATTCTAGAACTCCCATATGCTG	840							
Db	791	AGCTAAACATTGGATACATAGAACCTTAAGGCTCAGATTCTAGAACTCCCATATGCTG	850							
Qy	841	GAGATGTTAGCATGTTCTTTGTTGCTTCCAGATGAAATTTGCCGATGTGTCCACTGGCTGG	900							
Db	851	GAGATGTTAGCATGTTCTTTGTTGCTTCCAGATGAAATTTGCCGATGTGTCCACTGGCTGG	910							
Qy	901	AGCTCTGGAAGTGAATAAATCTATGACAACTCAACAGTGGACCCAGCAAGACAAAA	960							
Db	911	AGCTCTGGAAGTGAATAAATCTATGACAACTCAACAGTGGACCCAGCAAGACAAAA	970							
Qy	961	TGGCTGCAAGATGAAGTTGAGGTATACATACCCAGTTTCAATTTAGAGAGCATTTATGAC	1020							
Db	971	TGGCTGCAAGATGAAGTTGAGGTATACATACCCAGTTTCAATTTAGAGAGCATTTATGAC	1030							
Qy	1021	TCAGATCCATTCTGAAAGCATGGGATGGAGACGCTTCAACAGAGGAGGGCCCAATT	1080							
Db	1031	TCAGATCCATTCTGAAAGCATGGGATGGAGACGCTTCAACAGAGGAGGGCCCAATT	1090							

QY 1081 TCTCAGGATGTCGAGAGGAATGACCTGTTCTTTCTGAAGTGTCCACCAAGCCATGG 1140
 Db 1091 TCTCAGGATGTCGAGAGGAATGACCTGTTCTTTCTGAAGTGTCCACCAAGCCATGG 1150
 QY 1141 TGGATGTCGAGAGGACCTGAGCAGGCGCTGGCAGGAGGTGTTATCAGAGGA 1200
 Db 1151 TGGATGTCGAGAGGACCTGAGCAGGCGCTGGCAGGAGGTGTTATCAGAGGA 1210
 QY 1201 GAACCTGGACATGGAGGCCACACAGTTGTGGCAGATCATCGTCTTTCTTATATGC 1260
 Db 1211 GAACCTGGACATGGAGGCCACACAGTTGTGGCAGATCATCGTCTTTCTTATATGC 1270
 QY 1261 ATAAGATACCAAGTGCATTTATTTTTCGGCAGATTTCTCACCCTAAACCTAAGCGT 1320
 Db 1271 ATAAGATACCAAGTGCATTTATTTTTCGGCAGATTTCTCACCCTAAACCTAAGCGT 1330
 QY 1321 GCTGCTTCTGAAAAGATTTTCTAGATGAGCTGTGCGCTCAGAAATGCTATTTCAAAT 1380
 Db 1331 GCTGCTTCTGAAAAGATTTTCTAGATGAGCTGTGCGCTCAGAAATGCTATTTCAAAT 1390
 QY 1381 TGCCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAACCTTCGTACC 1440
 Db 1391 TGCCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAACCTTCGTACC 1450
 QY 1441 CACTAAATAAAACACAGAAATATTTAGCAATTTCTATATTAACATGCAACCCATT 1500
 Db 1451 CACTAAATAAAACACAGAAATATTTAGCAATTTCTATATTAACATGCAACCCATT 1510
 QY 1501 AATCAATTTGGTCTTCTAAAATGGGATCATGCCATTTAGATTTTCTTACTATCAGTTTA 1560
 Db 1511 AATCAATTTGGTCTTCTAAAATGGGATCATGCCATTTAGATTTTCTTACTATCAGTTTA 1570
 QY 1561 TTTTATACATTAACCTTTTACTTTGTTATTTATTTATATTAATGCTGAGTTTAA 1620
 Db 1571 TTTTATACATTAACCTTTTACTTTGTTATTTATTTATATTAATGCTGAGTTTAA 1630
 QY 1621 ATTATGCTCAGTCCCTTATTTAATGTAGCTAATAAAGTTATAGAGCGAGATGATCTGTTA 1680
 Db 1631 ATTATGCTCAGTCCCTTATTTAATGTAGCTAATAAAGTTATAGAGCGAGATGATCTGTTA 1690
 QY 1681 ATTCTCTATCTAATAAGCTTTAATTTGTTCTCATATGAAGTAATAGTAGGTATCCCT 1740
 Db 1691 ATTCTCTATCTAATAAGCTTTAATTTGTTCTCATATGAAGTAATAGTAGGTATCCCT 1749
 QY 1741 CCATGCCCTTCTGTAATAATATCTGGAATAAATTAAGCAATAGGCAATATATGTTA 1800
 Db 1750 CCATGCCCTTCTGTAATAATATCTGGAATAAATTAAGCAATAGGCAATATATGTTA 1809
 QY 1801 TGTGCAATTTCTAGAAATACATAACATATATGTCGTATCTTATATTCATATTCGAAG 1860
 Db 1810 TGTGCAATTTCTAGAAATACATAACATATATGTCGTATCTTATATTCATATTCGAAG 1869
 QY 1861 TATATAATAATAAACCTGCTTCCAAAC 1891
 Db 1870 TATATAATAATAAACCTGCTTCCAAAC 1900

RESULT 6

AA070474 standard; cDNA; 2409 BP.

ID AA070474

XX AA070474;

XX 25-MAR-2003 (revised)

DT 13-MAY-1991 (first entry)

XX Sequence encoding human minactivin.

XX Tumour location; tumour inhibition; chronic inflammation;

XX rheumatoid arthritis; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 PH CDS 49..1296
 FT /*tag= a
 XX EP238275-A.
 XX 23-SEP-1987.
 XX 13-MAR-1987; 87EP-00302200.
 XX 13-MAR-1986; 86AU-00005017.
 PR 17-MAY-1986; 87AU-00071855.
 PR 22-MAY-1986; 86AU-00006033.
 PR 18-SEP-1986; 86AU-00008100.
 PR 21-NOV-1986; 86AU-00009104.
 XX (BIOT-) BIOTECHN AUSTR PTY.
 PA (AUSU) AUSTR NAT UNIV.
 PA Antalis TM, Barnes TM, Clark MA, Devine PL, Goss NH, Lehrbach PR;
 PI WPI; 1987-265615/38.
 DR P-PSDB; AAP70300.
 XX DNA sequence encoding minactivin - used for monitoring and treating
 PT tumours and chronic inflammation such as rheumatoid arthritis.
 PS Claim 3; pp34-37; 76pp; English.
 XX Minactivin is produced by induced human monocytes, certain macrophages
 CC and transformed cells of monocytic lineage (see WO 86/01212). Using the
 CC AA sequence obtd. for peptides of minactivin, oligonucleotide probes can
 CC be synthesised. (see AAN70473, and AAN71181-N71183). The
 CC specific oligonucleotide probe may be radiolabelled and then used to
 CC screen cDNA libraries to identify clones contg. all or part of the
 CC minactivin gene. (Updated on 25-MAR-2003 to correct PI field.)
 XX Sequence 2409 BP; 743 A; 500 C; 508 G; 658 T; 0 U; 0 Other;

Query Match 96.8%; Score 1846; DB 1; Length 2409;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1849; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 15 GTCCAGACAGCACTCAGAGAAATACAGAGAAACACAGAGATTTGAAACAAATCGAGATCTT 74
 Db 1 GTCCAGACAGCACTCAGAGAAATACAGAGAAACACAGAGATTTGAAACAAATCGAGATCTT 60
 QY 75 TGTGTGGGAAACACACTCTTTGCCCTCAATTTATTCAGCATCTGGCAAAAGCAAGCCCC 134
 Db 61 TGTGTGGGAAACACACTCTTTGCCCTCAATTTATTCAGCATCTGGCAAAAGCAAGCCCC 120
 QY 135 ACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGTCTACATG 194
 Db 121 ACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGTCTACATG 180
 QY 195 GGCTCCAGGGGAGCAGCAGCAAGCCAGATGGCCAGGCTTCAGTTTAAATGAAGTGGGA 254
 Db 181 GGCTCCAGGGGAGCAGCAGCAAGCCAGATGGCCAGGCTTCAGTTTAAATGAAGTGGGA 240
 QY 255 GCCAATGCAGTTACCCCCCATGACTCCAGAGAACTTTACAGCTGTGGGTTCATGCAGCAG 314
 Db 241 GCCAATGCAGTTACCCCCCATGACTCCAGAGAACTTTACAGCTGTGGGTTCATGCAGCAG 300
 QY 315 ATCCGAGAGGGTAGTTATCTCTGATGCGATTTTGAGGACACAGCTGCAGATATAATCCAT 374
 Db 301 ATCCGAGAGGGTAGTTATCTCTGATGCGATTTTGAGGACACAGCTGCAGATATAATCCAT 360
 QY 375 TCATCCCTTCGGCTCTCTCAGCTCTGCAATCAATGATCCACAGGAAATTTTACTGGAA 434
 Db 361 TCATCCCTTCGGCTCTCTCAGCTCTGCAATCAATGATCCACAGGAAATTTTACTGGAA 420
 QY 435 AGTGTCAATAAGCTGTTTGGTGAGAGTCTCGAGCTTCCGGGAAAGATATATTCGACTC 494

Db 421 ACTGTCATTAAGCTGTTTGGTGAGAGTCTGCGAGCTTCGGGAGAGATATATTCGACTC 480
 Qy 495 TGTGAGAAATATATCTCTCAGAACCCGAGGAGTAGACTTCCTAGAGATGCGAGAGAA 554
 Db 481 TGTGAGAAATATATCTCTCAGAACCCGAGGAGTAGACTTCCTAGAGATGCGAGAGAA 540
 Qy 555 GCTAGAGAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTTG 614
 Db 541 GCTAGAGAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTTG 600
 Qy 615 TTACTGAGAGTCTGTAGATGGGATACAGAGATGCTGCTGGTGAATGCTGTCTACTTC 674
 Db 601 TTACTGAGAGTCTGTAGATGGGATACAGAGATGCTGCTGGTGAATGCTGTCTACTTC 660
 Qy 675 AAAGGAAAGTGGAAACTCCATTTGAGAGAAACTAAATGGGCTTTATPCCTTTCCGTGTA 734
 Db 661 AAAGGAAAGTGGAAACTCCATTTGAGAGAAACTAAATGGGCTTTATPCCTTTCCGTGTA 720
 Qy 735 AACTGGCTCAGCGCACACCTGTACAGATGATGCTGGTGAAGAGCTAAACATTGGA 794
 Db 721 AACTGGCTCAGCGCACACCTGTACAGATGATGCTGGTGAAGAGCTAAACATTGGA 780
 Qy 795 TACATAGAGACCTAAAGGCTCAGATTCAGAACTCCCATATGCTGGAGATGTTAGCATG 854
 Db 781 TACATAGAGACCTAAAGGCTCAGATTCAGAACTCCCATATGCTGGAGATGTTAGCATG 840
 Qy 855 TTCTTCTGCTTCCAGATGAAATTTGCCGATGTGTCCACTGGCTTGGAGCTGTGGAAGT 914
 Db 841 TTCTTCTGCTTCCAGATGAAATTTGCCGATGTGTCCACTGGCTTGGAGCTGTGGAAGT 900
 Qy 915 GAAATACCTATGACAACTCAACAGTGGACCGAGAAAGACAAATGGCTGGAAGATGA 974
 Db 901 GAAATACCTATGACAACTCAACAGTGGACCGAGAAAGACAAATGGCTGGAAGATGA 960
 Qy 975 GTTGAGGTATACATACCCAGTTCAAATTTAGAGAGCAATTAAGACTCAGATCCATTCTG 1034
 Db 961 GTTGAGGTATACATACCCAGTTCAAATTTAGAGAGCAATTAAGACTCAGATCCATTCTG 1020
 Qy 1035 AAAGCATGGGATGGAGGAGCCCTTCAAGAGGACGGGCAATTTCTCAGGATGTCG 1094
 Db 1021 AGAAGCATGGGATGGAGGAGCCCTTCAAGAGGACGGGCAATTTCTCAGGATGTCG 1080
 Qy 1095 GAGAGGAATGACCTGTTCTTCTGAGGTTCACCAAGGCAATGCTGGATGTAATGAG 1154
 Db 1081 GAGAGGAATGACCTGTTCTTCTGAGGTTCACCAAGGCAATGCTGGATGTAATGAG 1140
 Qy 1155 GAGGCACTGAGCAGCGCTGGACAGAGGTGTATGACAGGAGAGCTGCAATGGA 1214
 Db 1141 GAGGCACTGAGCAGCGCTGGACAGAGGTGTATGACAGGAGAGCTGCAATGGA 1200
 Qy 1215 GGCCACAGTTTGTGGCAGATCATCCGTTTCTTTTCTTATTTATGATAGATAACCAAG 1274
 Db 1201 GGCCACAGTTTGTGGCAGATCATCCCTTTTCTTTTCTTATTTATGATAGATAACCAAC 1260
 Qy 1275 TGCATTTATTTTTCGGCAGATTTTCTCACCCTAACTAAGCTGCTCTGCAAA 1334
 Db 1261 TGCATTTATTTTTCGGCAGATTTTCTCACCCTAACTAAGCTGCTCTGCAAA 1320
 Qy 1335 AGATTTTGTAGATGAGCTGTGTGCTCAGAAATGCTATTTCAAAATGCAAAATTTAG 1394
 Db 1321 AGATTTTGTAGATGAGCTGTGTGCTCAGAAATGCTATTTCAAAATGCAAAATTTAG 1380
 Qy 1395 AGATGTTTCTACATATTTCTGCTTCTGCAACCTTCTGCTACCCACTAAATAAAAC 1454
 Db 1381 AGATGTTTCTACATATTTCTGCTTCTGCAACCTTCTGCTACCCACTAAATAAAAC 1440
 Qy 1455 ACAGAAATTAATAGCAATTTGCTATTATPAAATGACAAACCTTAATTAATTTGCTTT 1514
 Db 1441 ACAGAAATTAATAGCAATTTGCTATTATPAAATGACAAACCTTAATTAATTTGCTTT 1500
 Qy 1515 CTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTATTTTATAACATTA 1574

Db 1501 CTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTATTTTATAACATTA 1560
 Qy 1575 ACTTTTACTTTCTTATTTTATTTATTTATATAAATGGTGAGTTTAAATTTATTTGCTCACTG 1634
 Db 1561 ACTTTTACTTTCTTATTTTATTTATTTATATAAATGGTGAGTTTAAATTTATTTGCTCACTG 1620
 Qy 1635 CTTATTTATGATGCTAATAAGTTATAGAGATGATCTGTTAAATTTCTTATCTCTAAT 1694
 Db 1621 CTTATTTATGATGCTAATAAGTTATAGAGATGATCTGTTAAATTTCTTATCTCTAAT 1680
 Qy 1695 AAATGCTTTTAAATTTCTCATTAATGAAGATTAAGTAGGTATCCCTCCATCCCTTCTGT 1754
 Db 1681 AAATGCTTTTAAATTTCTCATTAATGAAGATTAAGTAGGTATCCCTCCATCCCTTCTGT 1740
 Qy 1755 AATAATATCTGGAAACAAATTAACATAGCAATATATGTTATGTCATTTCTAGA 1814
 Db 1741 AATAATATCTGGAAACAAATTAACATAGCAATATATGTTATGTCATTTCTAGA 1800
 Qy 1815 AATACATAACACATATATATGCTGTATCTTATTTATTTCAATTCGAAGTATATAAT 1869
 Db 1801 AATACATAACACATATATATGCTGTATCTTATTTATTTCAATTCGAAGTATATAAT 1854

RESULT 7
 AAN91177
 ID AAN91177 standard; DNA; 2424 BP.
 XX
 AC AAN91177;
 XX
 DT 25-MAR-2003 (revised)
 DT 10-MAR-2003 (revised)
 DT 12-JUN-1989 (first entry)
 XX
 DE Sequence encoding plasminogen activator inhibitor, type 2 (PAI-2).
 XX
 KW Plasminogen activator inhibitor, type 2; antibodies to PAI-2; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 553..1797
 FT polyA_signal /tag= a
 FT 2382..2424 /tag= b
 FT
 XX
 PN DE3722673-A.
 XX
 PD 19-JAN-1989.
 XX
 PF 09-JUL-1987; 87DE-03722673.
 XX
 PR 09-JUL-1987; 87DE-00713272.
 XX
 PA (BEHW) BEHRINGWERKE AG.
 XX
 PI Tor N, Lecander I, Astedt B;
 XX
 DR WPI; 1989-024876/04.
 DR P-PSDB; AAP94160.
 XX
 PT New DNA encoding complete plasminogen activator inhibitor type 2 -
 PT expressed protein and derived antibodies, useful therapeutically and
 PT diagnostically.
 XX
 PS Disclosure; Page 3-4; 4pp; German.
 XX
 CC Sequence encodes plasminogen activator inhibitor type 2 (PAI-2) which can
 CC be produced using recombinant DNA techniques. Pure PAI-2 can be produced
 CC in this way and has therapeutic uses. Antibodies to this protein can be
 CC used to diagnose PAI-2 genetic defects. Patent no. DE-3713272 described
 CC the 361 N-terminal AA's of this protein which are used here as probes to
 CC rescreen the original cDNA bank, allowing the new sequence, encoding the
 CC entire PAI-2 mol.to be isolated. This PAI differs from human placental

CC	PAI at posns. 120 (Asn replacing Asp), 404 (Asn replacing Lys), and 414 (Ser replacing Cys). In addition Proline at posn. 393 is encoded by CCT instead of CCG. (Updated on 10-MAR-2003 to add missing OS field.)	
CC	(Updated on 25-MAR-2003 to correct FI field.)	
XX		
SQ	Sequence 2424 BP; 723 A; 516 C; 533 G; 652 T; 0 U; 0 Other;	
	Query Match 91.7%; Score 1750; DB 1; Length 2424;	
	Best Local Similarity 96.6%; Pred. No. 0;	
	Matches 1823; Conservative 0; Mismatches 5; Indels 60; Gaps 1;	
QY	19 GACAGCACTCAGAGATAACACAGACACACAGATGGAACATGAGGATCTTTG 78	
DB	569 GACAGCACTCAGAGATAACACAGACACACAGATGGAACATGAGGATCTTTG 628	
QY	79 TGGCAACACACTCTTTGCCCTCAATTTATCAAGCATCTGGCAAAAGCAAGCCACCC 138	
DB	629 TGGCAACACACTCTTTGCCCTCAATTTATCAAGCATCTGGCAAAAGCAAGCCACCC 688	
QY	139 AGAACCTCTTCTCCCTCCCATGGAGCATCTCGTCCACCATGGCCATGGTCTACATGGGCT 198	
DB	689 AGAACCTCTTCTCCCTCCCATGGAGCATCTCGTCCACCATGGCCATGGTCTACATGGGCT 748	
QY	199 CCAGGGCAGCACCGAAGACAGATGGCAAGTGCTTCAGTTTAAATGAAGTGGAGCCA 258	
DB	749 CCAGGGCAGCACCGAAGACAGATGGCAAGTGCTTCAGTTTAAATGAAGTGGAGCCA 808	
QY	259 ATGCAGTTACCCCATGACTCCAGAGAACTTTACAGCTGTGGTTCATGCAGCAGATCC 318	
DB	809 ATGCAGTTACCCCATGACTCCAGAGAACTTTACAGCTGTGGTTCATGCAGCAGATCC 868	
QY	319 AGAAGGTAGTTATCTCGATGGATTTGCGGCAAGCTCCAGATAAATCCATTCAT 378	
DB	869 AGAAGGTAGTTATCTCGATGGATTTGCGGCAAGCTCCAGATAAATCCATTCAT 928	
QY	379 CTTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAAATTTACTGGAAAGTG 438	
DB	929 CTTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAAATTTACTGGAAAGTG 988	
QY	439 TCATAAGCTTTTGGTGAAGCTCGGAGCTTCGGGAAGAAATATATTCGACTCTGTC 498	
DB	989 TCAATAAGCTTTTGGTGAAGCTCTGGAGC----- 1020	
QY	499 AGAAATATTACTCTCAGAACCCCGGAGTAGCTTCTAGATGTGCAGAGAAGCTTA 558	
DB	1021 -----GACTTCTTAGAATGTGCAGAGAAGCTTA 1048	
QY	559 GAAABAAGATTAAATCTCGGTCAAGACTCAAAACCAAGGCAAAATCCCAACTTGTTC 618	
DB	1049 GAAABAAGATTAAATCTCGGTCAAGACTCAAAACCAAGGCAAAATCCCAACTTGTTC 1108	
QY	619 CTGAAGGTTCTGTAGATGGGATACAGGATGGTCTGGTGAATGTGTCTACTTCAAG 678	
DB	1109 CTGAAGGTTCTGTAGATGGGATACAGGATGGTCTGGTGAATGTGTCTACTTCAAG 1168	
QY	679 GAAAGTGAARAACCTCAATTTGAGAGAAACTAAATGGGCTTTATCTCTTCGGTGTAACT 738	
DB	1169 GAAAGTGAARAACCTCAATTTGAGAGAAACTAAATGGGCTTTATCTCTTCGGTGTAACT 1228	
QY	739 CGGCTCAGGCGCACCTGTGTACAGATGATGTCTGGTGAAGCTTAAACATTTGGATACA 798	
DB	1229 CGGCTCAGGCGCACCTGTGTACAGATGATGTCTGGTGAAGCTTAAACATTTGGATACA 1288	
QY	799 TAGAAGACCTTAAGGCTCAGATTTCTAGAACTCCCATATGCTGGAGATGTTAGCATGTTCT 858	
DB	1289 TAGAAGACCTTAAGGCTCAGATTTCTAGAACTCCCATATGCTGGAGATGTTAGCATGTTCT 1348	
QY	859 TGTTCCTCCAGATGAATTCGGATGTCTCCACTGGCTGGAGCTGCTCGAAAGTGA 918	
DB	1349 TGTTCCTCCAGATGAATTCGGATGTCTCCACTGGCTGGAGCTGCTCGAAAGTGA 1408	
QY	919 TAACTTATGACAACTCAACAGTGGACCGCAAGACAAATCGCTGGAAGTGAAGTTG 978	

RESULT 8
AAN80060
ID AAN80060 standard; cDNA; 1675 BP.
XX
AC AAN80060;
XX

DB	1409 TAACCTATGACAACTCAACAGCTGGACCAAGACAAATAGCTGAAGATGAAGTTG 1468	
QY	979 AGGTATACATACCCAGTTCAAATTAAGAGAGATTAAGAACTCAGATCCATTTCTGA 1038	
DB	1469 AGGTATACATACCCAGTTCAAATTAAGAGAGATTAAGAACTCAGATCCATTTCTGA 1528	
QY	1039 GATGGGCATGAGGAGCGCTTCAACAGGAGCGGCGCAATTTCTCAGGATGTCGGAGA 1098	
DB	1529 GATGGGCATGAGGAGCGCTTCAACAGGAGCGGCGCAATTTCTCAGGATGTCGGAGA 1588	
QY	1099 GGAATGACCTGTTTCTTTCTGAAAGTGTTCACCAAGCCATGGTGGATGTAATGAGG 1158	
DB	1589 GGAATGACCTGTTTCTTTCTGAAAGTGTTCACCAAGCCATGGTGGATGTAATGAGG 1648	
QY	1159 GCACTGAAGCAGCGCTGGCAGAGAGGTGTATGACAGGAGAGAACTCGACATGAGGCG 1218	
DB	1649 GCACTGAAGCAGCGCTGGCAGAGAGGTGTATGACAGGAGAGAACTCGACATGAGGCG 1708	
QY	1219 CACAGTTTGGCAGATCATCGTTTCTTTTCTTATTTATGATGAATGAATGAATGA 1278	
DB	1709 CACAGTTTGGCAGATCATCGTTTCTTTTCTTATTTATGATGAATGAATGAATGA 1768	
QY	1279 TTTTATTTTGGCAGATTTTCTCACTCACTCACTCACTCACTCACTCACTCACTCA 1338	
DB	1769 TTTTATTTTGGCAGATTTTCTCACTCACTCACTCACTCACTCACTCACTCACTCA 1828	
QY	1339 TTTTGTAGATGAGCTGTGCTCAGAAATGCTTATTTCAATTTGCCAAATTTTGA 1398	
DB	1829 TTTTGTAGATGAGCTGTGCTCAGAAATGCTTATTTCAATTTGCCAAATTTTGA 1888	
QY	1399 GTTTTCTACATATTTCTGCTCTCTGAACTCTGCTACCCACTCACTCACTCACTCA 1458	
DB	1889 GTTTTCTACATATTTCTGCTCTCTGAACTCTGCTACCCACTCACTCACTCACTCA 1948	
QY	1459 AATAATTTAGACATTTCTGCTTATTAATGATGATGATGATGATGATGATGATGAT 1518	
DB	1949 AATAATTTAGACATTTCTGCTTATTAATGATGATGATGATGATGATGATGATGAT 2008	
QY	1519 AATGGGATCATGCCCATTTAGATTTCTTACTATCATGATTTATTTTATAAATTA 1578	
DB	2009 AATGGGATCATGCCCATTTAGATTTCTTACTATCATGATTTATTTTATAAATTA 2068	
QY	1579 TTACTTTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 1638	
DB	2069 TTACTTTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 2128	
QY	1639 TTTAATGTAGCTAATAAAGTTATGAGACAGATGATCTGTTAATTTCTTATCTTA 1698	
DB	2129 TTTAATGTAGCTAATAAAGTTATGAGACAGATGATCTGTTAATTTCTTATCTTA 2188	
QY	1699 GCCTTTAATTTCTCTAATAATGAAGATAAGTAGGTATCCCTCCATGCCCTTCTG 1758	
DB	2189 GCCTTTAATTTCTCTAATAATGAAGATAAGTAGGTATCCCTCCATGCCCTTCTG 2248	
QY	1759 AATATCTGGAAAAACATTAACATAGGCAATATATGTTATGTCATTTCTAGAAATA 1818	
DB	2249 AATATCTGGAAAAACATTAACATAGGCAATATATGTTATGTCATTTCTAGAAATA 2308	
QY	1819 CATAACACATATATGCTGTGATCTTATTTCAATTTGAAGATGATATAATAATAACCT 1878	
DB	2309 CATAACACATATATGCTGTGATCTTATTTCAATTTGAAGATGATATAATAATAACCT 2368	
QY	1879 GCTTCCCAACCAACAAAAAATAAAAAA 1906	
DB	2369 GCTTCCCAACCAACAAAAAATAAAAAA 2396	

05-NOV-1990 (first entry)
 Sequence encoding plasminogen activator inhibitor type 2 (PAI2) in human
 placental cDNA clone lambda-plc 1.
 Protease inhibitor; thrombolytic; ss.
 Homo sapiens.
 Key Location/Qualifiers
 CDS 1..1086
 /*tag= a
 DE3713272-A.
 03-NOV-1988.
 18-APR-1987; 87DE-03713272.
 09-JUL-1987; 87DE-03722673.
 (BEHW) BEHRINGWERKE AG.
 Ny T, Lecander I, Astedt B;
 WPI; 1988-315694/45.
 P-PSDB; AAP80058.
 New DNA sequence encoding plasminogen activator inhibitor type 2 - and
 derived proteins and antibodies, useful therapeutically and in diagnosis.
 Claim 1; Table 1 Page 5; 6pp; German.
 Commercially available human placental cDNA bank in lambda gtl1 phage was
 used to transform E.coli Y1090 and culture supernatants from the
 transfectants screened by reaction with mouse Ab. against PAI2. Clone
 lambda-plc 1 was one of 6 clones isolated. PAI2 is a strong and specific
 inhibitor for the active forms of PA so is used to treat or prevent
 disorders associated with abnormal levels of PA. PA is implicated in
 haemostasis, tissue repair, metastasis formation, ovulation, fertility,
 macrophage migration and inflammatory processes
 Sequence 1675 BP; 530 A; 322 C; 331 G; 492 T; 0 U; 0 Other;
 Query Match 87.4%; Score 1667; DB 1; Length 1675;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1670; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 225 GCCAAGGTCCTCAGTTAATGAAGTGGAGCCCAATGACGTTACCCCATGACTCCAGAG 284
 DB 1 GCCAAGGTCCTCAGTTAATGAAGTGGAGCCCAATGACGTTACCCCATGACTCCAGAG 60
 QY 285 AACTTTACCAGCTGTGGGTTTCATGCAGCAGATCCAGAAGGGTAGTTATCCTGATGCGATT 344
 DB 61 AACTTTACCAGCTGTGGGTTTCATGCAGCAGATCCAGAAGGGTAGTTATCCTGATGCGATT 120
 QY 345 TTGCAGGCAACAGCTGCAGTAAATCCATTCTCCTTCGCTCTCAGCTCTGCAATC 404
 DB 121 TTGCAGGCAACAGCTGCAGTAAATCCATTCTCCTTCGCTCTCAGCTCTGCAATC 180
 QY 405 AATGCAATCCACAGGGAATTTATTTACTGGAAGTGTCAATGAAGCTGTTGGTGAAGTCT 464
 DB 181 AATGCAATCCACAGGGAATTTATTTACTGGAAGTGTCAATGAAGCTGTTGGTGAAGTCT 240
 QY 465 GCGAGCTTCGGAAGGAATATTTCTGCTGTGAGAAATATTACTCTCAGAACCCCGAG 524
 DB 241 GCGAGCTTCGGAAGGAATATTTCTGCTGTGAGAAATATTACTCTCAGAACCCCGAG 300
 QY 525 GCAGTAGACTTCTAGAAATGTGAGAAAGAGCTAGAAAAAGATTAATTCCTGGGTCAAG 584
 DB 301 GCAGTAGACTTCTAGAAATGTGAGAAAGAGCTAGAAAAAGATTAATTCCTGGGTCAAG 360
 QY 585 ACTCAACCAAGGCAAAATCCCAAACTTGTTACCTGAAGTTCTGTAGATGGGATACC 644

Db 361 ACTCAACCAAGGCAAAATCCCAAACTTGTTACTGAAGTTCTGTAGATGGGATACC 420
 QY 645 AGGATGGTCTGGTGAATGCTGTCTACTTCAAGGAAAGTGGAAACTCCATTGAGAAG 704
 Db 421 AGGATGGTCTGGTGAATGCTGTCTACTTCAAGGAAAGTGGAAACTCCATTGAGAAG 480
 QY 705 AAACCTAAATGGGCTTTATCCCTTTCGGTGTAAACTCGGCTCAGCGCACACCTGTACAGATG 764
 Db 481 AAACCTAAATGGGCTTTATCCCTTTCGGTGTAAACTCGGCTCAGCGCACACCTGTACAGATG 540
 QY 765 ATGTACTTGGGTGAAAGCTAAACATTTGGATATCATAGAACCTTAAGGCTCAGATTCTA 824
 Db 541 ATGTACTTGGGTGAAAGCTAAACATTTGGATATCATAGAACCTTAAGGCTCAGATTCTA 600
 QY 825 GAACTCCCATATGCTGGAGATGTAGCATGTTCTTGTGCTCCAGATGAAATTCGCGAT 884
 Db 601 GAACTCCCATATGCTGGAGATGTAGCATGTTCTTGTGCTCCAGATGAAATTCGCGAT 660
 QY 885 GTGTCCACTGGCTTGGAGCTGTGGAAGTGAATTAACCTATGACAACTCAACAAGTGG 944
 Db 661 GTGTCCACTGGCTTGGAGCTGTGGAAGTGAATTAACCTATGACAACTCAACAAGTGG 720
 QY 945 ACCAGCAAGACAAAATGGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTA 1004
 Db 721 ACCAGCAAGACAAAATGGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTA 780
 QY 1005 GAAGAGCATTTATGAATCTAGATCCATTCTGAAAGCATGGGCATGAGGACGCTTCAAC 1064
 Db 781 GAAGAGCATTTATGAATCTAGATCCATTCTGAAAGCATGGGCATGAGGACGCTTCAAC 840
 QY 1065 AAGGACCGGCGCAATTTCTCAGGAGTGTGGAGGAAATGACCTGTTCTTCTTGAAGTG 1124
 Db 841 AAGGACCGGCGCAATTTCTCAGGAGTGTGGAGGAAATGACCTGTTCTTCTTGAAGTG 900
 QY 1125 TTCCACCAAGCCATGGTGAATGTGAATGAGGAGGCACTGAAGCAGCGCTGGCACAGGA 1184
 Db 901 TTCCACCAAGCCATGGTGAATGTGAATGAGGAGGCACTGAAGCAGCGCTGGCACAGGA 960
 QY 1185 GGTGTTATGACAGGAGAACTGGACATGGAGGCGCCACAGTTTGTGGCAGATCATCCGTTT 1244
 Db 961 GGTGTTATGACAGGAGAACTGGACATGGAGGCGCCACAGTTTGTGGCAGATCATCCCTTT 1020
 QY 1245 CTTTCTTCTTATGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 1304
 Db 1021 CTTTCTTCTTATGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 1080
 QY 1305 CCCTAAAACCTAAGCGTGTCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCGCTCAG 1364
 Db 1081 CCCTAAAACCTAAGCGTGTCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCGCTCAG 1140
 QY 1365 AATTGCTATTTCAAAATGGCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTCTG 1424
 Db 1141 AATTGCTATTTCAAAATGGCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTCTG 1200
 QY 1425 AACAACTTCTGCTACCCACTAAATAAAACACAGAAATAAATAGCAATTTCTATTATA 1484
 Db 1201 AACAACTTCTGCTACCCACTAAATAAAACACAGAAATAAATAGCAATTTCTATTATA 1260
 QY 1485 ACATGACAACTTAAATCAATTTGGTCTTCTAAATGGGATCATGCCATTTAGATTTT 1544
 Db 1261 ACATGACAACTTAAATCAATTTGGTCTTCTAAATGGGATCATGCCATTTAGATTTT 1320
 QY 1545 CCTTACTATCAGTTATTTTATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1604
 Db 1321 CCTTACTATCAGTTATTTTATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
 QY 1605 AATGGTGAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1664
 Db 1381 AATGGTGAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440
 QY 1665 AGCAGATGATCTGTTAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1724

QY 1082 CTCAGGATGTCGGAGAGGATGACCTGTTCTTTCTGAAGTGTTCACCAAGCCATGGT 1141
DB 972 CTCAGGATGTCGGAGAGGATGACCTGTTCTTTCTGAAGTGTTCACCAAGCCATGGT 1031
QY 1142 GGAATGTAATGAGAGGGCACTGAAGCAGCGCGCTGGGCACAGAGGTGTATGACGGAG 1201
DB 1032 GGATGTAATGAGAGGGCACTGAAGCAGCGCGCTGGGCACAGAGGTGTATGACGGAG 1091
QY 1202 AACTGACATGGAGGCCACAGATTGTGGCAGATCATCCGTTTCTTTCTTATTATGCA 1261
DB 1092 AACTGACATGGAGGCCACAGATTGTGGCAGATCATCCCTTTCTTTCTTATTATGCA 1151
QY 1262 TAGATTAACCAAGTGCATTTATTTTTCGGCAGATTTGCTCACCTAAACTRAGCGTG 1321
DB 1152 TAGATTAACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCTAAACTRAGCGTG 1211
QY 1322 CTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAGAAATGCTATTTCAAAT 1381
DB 1212 CTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAGAAATGCTATTTCAAAT 1271
QY 1382 GCCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAACAACCTTCTGTACCC 1441
DB 1272 GCCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAACAACCTTCTGTACCC 1331
QY 1442 ACTAAATPAAAAACACAGAAATAATTAGACAATGTCTATTATTAACATGACAAACCTATT 1501
DB 1332 ACTAAATPAAAAACACAGAAATAATTAGACAATGTCTATTATTAACATGACAAACCTATT 1391
QY 1502 ATCAATTTGGTCTTCTAAATGGGATCATGCCCAATTTAGATTTTCTTACTATCAGTTTAT 1561
DB 1392 ATCAATTTGGTCTTCTAAATGGGATCATGCCCAATTTAGATTTTCTTACTATCAGTTTAT 1451
QY 1562 TTTTATAACATTAATTTTACTTTGTATTATTTATTTATTTATTAATGGTGGT 1618
DB 1452 TTTTATAACATTAATTTTACTTTGTATTATTTATTTATTTATTAATGGTGGT 1508

RESULT 11
AAQ11820
ID AAQ11820 standard; DNA; 1482 BP.
XX
AC AAQ11820;
XX
DT 09-JAN-2003 (revised)
DT 18-SEP-1991 (first entry)
XX
DE Plasminogen activator inhibitor-2 variant.
XX PAI-2; protease sensitive site; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 22..1167
FT FT /*tag= a
FT FT /product= "PAI-2 variant"
XX
PN WO9109124-A.
XX
PD 27-JUN-1991.
XX
PF 20-DEC-1989; 89AU-00007924.
XX
PR 20-DEC-1989; 89AU-00007924.
XX
PA (BIOT-) BIOTECH AUST PTY LT.
XX
XX Goss NH, Richardson MA;
XX
XX WPI; 1991-208150/28.
XX
XX P-P8DB; AAE12144.
XX

PT Plasminogen activator inhibitor-2 variants with modified protease - used
PT to treat inflammatory diseases e.g. rheumatoid arthritis, ulcerative
XX colitis and psoriasis and fibrinolytic diseases.
XX
PS Disclosure; Page 48-50; 88pp; English.
XX
CC The variant has amino acids 66-98 deleted. It retains the biological
CC activity while lacking protease sensitivity of native PAI-2. The variant
CC is prep'd. by alteration of the DNA encoding PAI-2 and expression in
CC hosts. DNA sequences encoding amino acids 66-98 were deleted from PAI-2
CC using the PCR technique of site-directed mutagenesis by overlap
CC extension. See also AAQ11619. (Updated on 09-JAN-2003 to add missing OS
CC field.)
XX
SQ Sequence 1482 BP; 438 A; 308 C; 320 G; 416 T; 0 U; 0 Other;
Query Match 70.3%; Score 1341; DB 2; Length 1482;
Best Local Similarity 93.3%; Pred. No. 7.4e-268; Indels 99; Gaps 1;
Matches 1453; Conservative 0; Mismatches 5;
QY 62 AATGGAGGATCTTTGTGTGCAACACACACTCTTTTGGCCCTCAATTTATTCAAGCATCTGGC 121
DB 21 AATGGAGGATCTTTGTGTGCAACACACACTCTTTTGGCCCTCAATTTATTCAAGCATCTGGC 80
QY 122 ABAAGCAAGCCCAACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGGC 181
DB 81 ABAAGCAAGCCCAACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGGC 140
QY 182 CATGGTCTCATGGGCTCCAGGGGAGCAGCAGAACAGATGGCAGAGTCTTCAGTT 241
DB 141 CATGGTCTCATGGGCTCCAGGGGAGCAGCAGAACAGATGGCAGAGTCTTCAGTT 200
QY 242 TAATGAAGTGGGAGCAATGAGTTACCCCATGACTCCAGAGAACTTTTACCAGCTGTGG 301
DB 201 TAATGAAGTGGGAGCC----- 216
QY 302 GTTCATCGACGAGATCCAGAGGGGTAGTTATCCCTGATGCGATTTTTCAGGACACAAGCTGC 361
DB 217 -----GCTGC 221
QY 362 AGATAAAATCCATTCATCCCTTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAA 421
DB 222 AGATAAAATCCATTCATCCCTTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAA 281
QY 422 TTATTTACTGGAAAGTGTCATTAAGCTGTTTGGTGAGAAGTCTGGAGCTTCCGGGAAGA 481
DB 282 TTATTTACTGGAAAGTGTCATTAAGCTGTTTGGTGAGAAGTCTGGAGCTTCCGGGAAGA 341
QY 482 ATATATTGCACTGTGTCAGAAATATTACTCTCTCAGAACCCAGGAGTAGACTTCTCTAGA 541
DB 342 ATATATTGCACTGTGTCAGAAATATTACTCTCTCAGAACCCAGGAGTAGACTTCTCTAGA 401
QY 542 ATGTGCAGAGAAGCTAGAAAAAGATTAAATTCCTGGGTCAAGACTCAAAACCAAGGCAA 601
DB 402 ATGTGCAGAGAAGCTAGAAAAAGATTAAATTCCTGGGTCAAGACTCAAAACCAAGGCAA 461
QY 602 AATCCCAACTGTGTACTGGAAGTCTGTAGATGGGATACAGGATGGTCTCTGGTGAA 661
DB 462 AATCCCAACTGTGTACTGGAAGTCTGTAGATGGGATACAGGATGGTCTCTGGTGAA 521
QY 662 TGCTGTCTACTTCAAAGGAAAGTGGAAAACTCCATTGTGAGAAGAACTAAATGGGCTTTA 721
DB 522 TGCTGTCTACTTCAAAGGAAAGTGGAAAACTCCATTGTGAGAAGAACTAAATGGGCTTTA 581
QY 722 TCCCTTCCGCTGTAAACTCGGCTCAGCGCACACTGTACAGATGATGTACTTGGGTGAAAA 781
DB 582 TCCCTTCCGCTGTAAACTCGGCTCAGCGCACACTGTACAGATGATGTACTTGGGTGAAAA 641
QY 782 GCTAAACATTTGATACATAGAGACCTAAAGGCTCAGATTCTAGAACTCCCATATGCTGG 841
DB 642 GCTAAACATTTGATACATAGAGACCTAAAGGCTCAGATTCTAGAACTCCCATATGCTGG 701
QY 842 AGATGTTAGCATGTTCTTTGTTGCTTCCAGATGAAATTTGCCGATGTGTCCACTGGCTTGG 901

Db 702 AGATGTTAGCATGTTCTTGTTGCTTCCAGATGAATGCCGATGTGTCACATGCTGCTGGA 761
Qy 902 GCTGCTGGAAGTGAATACCTATGACAACTCACAAGTGGACCAAGCAAAAT 961
Db 762 GCTGCTGGAAGTGAATACCTATGACAACTCACAAGTGGACCAAGCAAAAT 821
Qy 962 GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAAGAAGAGCATTAAGACT 1021
Db 822 GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAAGAAGAGCATTAAGACT 881
Qy 1022 CAGATCCATCTGAAAAGCATGGCATGGAGGCGCTTCAACAAGGAGCGGGCAATTT 1081
Db 882 CAGATCCATCTGAAAAGCATGGCATGGAGGCGCTTCAACAAGGAGCGGGCAATTT 941
Qy 1082 CTCAGGATGTCGAGAGGAGATGACCTGTTTCTTCTGAAGTGTCCACCAAGCATGCT 1141
Db 942 CTCAGGATGTCGAGAGGAGATGACCTGTTTCTTCTGAAGTGTCCACCAAGCATGCT 1001
Qy 1142 GGATGTGAATGAGGAGGCACTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGGAG 1201
Db 1002 GGATGTGAATGAGGAGGCACTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGGAG 1061
Qy 1202 AACTGGATGAGGCGCCACAGTTTGGCAGATCATCCGTTCTTCTTATTATGCA 1261
Db 1062 AACTGGATGAGGCGCCACAGTTTGGCAGATCATCCGTTCTTCTTATTATGCA 1121
Qy 1262 TAAGATAACCAAGTGCATTTATTTTCGCGAGATTTGCTCACCTAAAACCTAAGCGTG 1321
Db 1122 TAAGATAACCAAGTGCATTTATTTTCGCGAGATTTTCTCACCTAAAACCTAAGCGTG 1181
Qy 1322 CTGCTTCTGAAAAGATTTTGTAGATGAGTGTGCTCAGAAATGCTATTTCAAAT 1381
Db 1182 CTGCTTCTGAAAAGATTTTGTAGATGAGTGTGCTCAGAAATGCTATTTCAAAT 1241
Qy 1382 GCGAAAATTTAGAGATGTTTCTACATATTTCTGCTTCTGACAACTCTGCTACCC 1441
Db 1242 GCGAAAATTTAGAGATGTTTCTACATATTTCTGCTTCTGACAACTCTGCTACCC 1301
Qy 1442 ACTAAATAAAAAACACAGAAATAATATAGACAATTTGCTATTATTAACATGACAACCTATTA 1501
Db 1302 ACTAAATAAAAAACACAGAAATAATATAGACAATTTGCTATTATTAACATGACAACCTATTA 1361
Qy 1502 ATCATTTGCTCTTCTAAAAGGATCATGCCATTTAGATTTTCTTACTATCAGTTTAT 1561
Db 1362 ATCATTTGCTCTTCTAAAAGGATCATGCCATTTAGATTTTCTTACTATCAGTTTAT 1421
Qy 1562 TTTTATAACATTAACCTTTTACTTTGTTATTTATTTATTTATTAATGTTGAGTTT 1618
Db 1422 TTTTATAACATTAACCTTTTACTTTGTTATTTATTTATTTATTAATGTTGAGTTT 1478

RESULT 12
ID AAQ11128 standard; DNA; 1328 BP.

XX AAQ11128;
AC
XX
DT 09-JAN-2003 (revised)
DT 06-JUN-1991 (first entry)

XX DE PreA-plasminogen activator inhibitor-2 with signal sequence.
XX KW Plasminogen activator inhibitor; PAI-2; signal sequence; diagnosis;
XX KW treatment; tumour; inflammation; secretion; ss.

XX OS Synthetic.

XX FH Location/Qualifiers
XX CDS 15..1316
FT /tag= a
FT /product= "preB-PAI-2"
FT sig_peptide 15..71

FT FT /*tag= b
XX XX /note= "synthetic-claim3"
XX W09103556-A.
XX PD 21-MAR-1991.
XX PF 05-SEP-1989; 89AU-00006179.
XX PR 05-SEP-1989; 89AU-00006179.
XX PA (BIOT-) BIOTECHN AUST PTY L.
XX PI Whitfield P, Richardson MA, Bunn CL;
XX DR WPI: 1991-102074/14.
XX DR P-PSDB; AAR11352.
XX PT Prepn. of glycosylated plasminogen activator inhibitor-2 (PAI-2) - for
XX PT diagnosis and treatment of tumours, and in treatment of
XX PT immuno-suppressant conditions and inflammation.
XX PS Disclosure; Fig 7; 70pp; English.

XX CC Addition of a signal sequence to PAI-2 not only facilitates secretion of
XX CC the glycosylated 50 kD form of PAI-2 but also directs correct processing
XX CC of the signal. The signal-contg. variant was constructed such that the
XX CC artificial signal was fused to the NH2-terminal methionine of PAI-2.
XX CC Correct cellular processing of this signal peptide should result in a
XX CC mature PAI-2 molecule of 415 amino acids with Met at its NH2-terminus.
XX CC See AAQ11129 for PAI-2 with the alpha-1-antitrypsin signal sequence.
XX CC (Updated on 09-JAN-2003 to add missing OS field.)

XX SQ Sequence 1328 BP; 377 A; 300 C; 317 G; 334 T; 0 U; 0 Other;

Query Match 65.5%; Score 1249.4; DB 2; Length 1328;
Best Local Similarity 99.5%; Pred. No. 6.6e-249;
Matches 1253; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 61 CAATGAGAGATCTTTGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAGCATCTGG 120
Db 70 CTATGGAAGATCTTTGTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAGCATCTGG 129
Qy 121 CAAAGCAAGCCGCCACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACATGG 180
Db 130 CAAAGCAAGCCGCCACCCAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACATGG 189
Qy 181 CCATGCTCTACATGGGCTCCAGGGGAGACCCGAGACCGAGATGCCCAAGGTGCTTCAGT 240
Db 190 CCATGCTCTACATGGGCTCCAGGGGAGACCCGAGACCGAGATGCCCAAGGTGCTTCAGT 249
Qy 241 TTAATGAAGTGGAGCCCAATGCGAGTTACCCCATGACTCCAGAGAACTTTACCAAGCTGTG 300
Db 250 TTAATGAAGTGGAGCCCAATGCGAGTTACCCCATGACTCCAGAGAACTTTACCAAGCTGTG 309
Qy 301 GGTTTCATGACAGAGATCCAGAAAGGTAGTTATCTCTGATGCCATTTTGAGGCACCAAGCTG 360
Db 310 GGTTTCATGACAGAGATCCAGAAAGGTAGTTATCTCTGATGCCATTTTGAGGCACCAAGCTG 369
Qy 361 CAGATAAAATCCATTGATCTCTCGGCTCTCTCAGCTCTGCAATCAATCCATCCAGAGGA 420
Db 370 CAGATAAAATCCATTGATCTCTCGGCTCTCTCAGCTCTGCAATCAATCCATCCAGAGGA 429
Qy 421 ATTATTACTGGAAAGTGTCAATAAGCTGTTGTTGAGAAAGTCTCGAGAGTTCGGGGAAG 480
Db 430 ATTATTACTGGAAAGTGTCAATAAGCTGTTGTTGAGAAAGTCTCGAGAGTTCGGGGAAG 489
Qy 481 AATATATTCGACTCTGTGAGAAATATTACTCTCAGAACCCCGAGCAGTAGACTTCCTTAG 540
Db 490 AATATATTCGACTCTGTGAGAAATATTACTCTCAGAACCCCGAGCAGTAGACTTCCTTAG 549
Qy 541 AATGTGAGAAAGCTGTAGAAAAAGATTAATTTCTGGTCAAGACTCAAAACCAAGGCA 600

Db	550	AATGTCAGAGAGCTAGAAAAAGATTAAATCTCGGTCAAGACTCAAAACCAAGGCA	609
Qy	601	AAATCCCAACTGTTTACCTGAGGTTCTGTAGATGGGTACACAGATGCTCTGGTGA	660
Db	610	AAATCCCAACTGTTTACCTGAGGTTCTGTAGATGGGTACACAGATGCTCTGGTGA	669
Qy	661	ATGCTGCTACCTTCAAGGAAAGTGGAAACTCCATTTGAGAGAACTAAATGGGCTTT	720
Db	670	ATGCTGCTACCTTCAAGGAAAGTGGAAACTCCATTTGAGAGAACTAAATGGGCTTT	729
Qy	721	ATCCTTTCCGTTGAACTCCGCTCAGCGCACCTGTACAGATGATGTTGCTGAAA	780
Db	730	ATCCTTTCCGTTGAACTCCGCTCAGCGCACCTGTACAGATGATGTTGCTGAAA	789
Qy	781	AGCTAAACATGATACATAGAGACCTAAGGCTCAGATTTCTAGATCCCATATGCTG	840
Db	790	AGCTAAACATGATACATAGAGACCTAAGGCTCAGATTTCTAGATCCCATATGCTG	849
Qy	841	GAGATGTTAGCATGTTCTTGTGTTCCAGATGAAATGCGGATGTCCTGCTGG	900
Db	850	GAGATGTTAGCATGTTCTTGTGTTCCAGATGAAATGCGGATGTCCTGCTGG	909
Qy	901	AGCTGCTGGAAGTGAATTAACCTATGACAACTCAACAAGTGGACCAAGCAAA	960
Db	910	AGCTGCTGGAAGTGAATTAACCTATGACAACTCAACAAGTGGACCAAGCAAA	969
Qy	961	TGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAAATTAGAGCATTTATGAC	1020
Db	970	TGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAAATTAGAGCATTTATGAC	1029
Qy	1021	TCAGATCCATTTCTGAAAGCATGCGCATGGAGGAGCGCTTCAACAAGGAGCGGCAATT	1080
Db	1030	TCAGATCCATTTCTGAAAGCATGCGCATGGAGGAGCGCTTCAACAAGGAGCGGCAATT	1089
Qy	1081	TCTCAGGATGTCGGAGGAGTACACCTGTTCTTCTGAGTGTTCACCAAGCCATGG	1140
Db	1090	TCTCAGGATGTCGGAGGAGTACACCTGTTCTTCTGAGTGTTCACCAAGCCATGG	1149
Qy	1141	TGATGTGAATGAGAGGCGCACTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGA	1200
Db	1150	TGATGTGAATGAGAGGCGCACTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGA	1209
Qy	1201	GAACTGGACATGGAGGCGCCACAGTTTGGCAGATCATCCGTTCTTTCTATTATGC	1260
Db	1210	GAACTGGACATGGAGGCGCCACAGTTTGGCAGATCATCCGTTCTTTCTATTATGC	1269
Qy	1261	ATAAGATAACCAAGTGCATTTTTCGGCAGATTTTTCCTCACCCTAAACTAAGCG	1319
Db	1270	ATAAGATAACCAAGTGCATTTTTCGGCAGATTTTTCCTCACCCTAAACTAAGCG	1328
RESULT 13			
ID	AAQ11129 standard; DNA; 1340 BP.		
XX	AAQ11129;		
AC	AAQ11129;		
XX	(revised)		
DT	09-JAN-2003 (first entry)		
DT	06-JUN-1991		
XX	PreB-PAI-2 with alpha-1-antitrypsin signal sequence.		
DE	Plasminogen activator inhibitor; PAI-2; signal sequence; diagnosis;		
KW	treatment; tumour; inflammation; secretion; alpha-1-antitrypsin; ss.		
OS	Synthetic.		
XX	Location/Qualifiers		
PH	15..1328		
FT	/*tag= a		
FT	/*product= "preB-PAI-2"		
FT	15..86		
FT	/*tag= b		
FT	sig_peptide		
FT			

FT	XX	/label= alpha-1-antitrypsin
XX	PN	WO9103556-A.
XX	PD	21-MAR-1991.
XX	PF	05-SEP-1989; 89AU-00006179.
XX	PR	05-SEP-1989; 89AU-00006179.
XX	PA	(BIOT-) BIOTECHN AUST PTY L.
XX	PI	Whitfield P, Richardson MA, Bunn CL;
XX	PI	WPI; 1991-102074/14.
XX	DR	P-PSDB; AAR11353.
XX	PT	Prepn. of glycosylated plasminogen activator inhibitor-2 (PAI-2) - for
XX	PT	diagnosis and treatment of tumours, and in treatment of
XX	PT	immuno:suppressant conditions and inflammation.
XX	PS	Disclosure; Fig 8; 70pp; English.
XX	CC	Addition of a signal sequence to PAI-2 not only facilitates secretion of
XX	CC	the glycosylated 60 kD form of PAI-2 but also directs correct processing
XX	CC	of the signal. The signal-contg. variant was constructed such that the
XX	CC	human alpha-1-antitrypsin signal was fused to PAI-2 without its NH2-
XX	CC	terminal methionine. Correct cellular processing of this signal peptide
XX	CC	should result in a mature PAI-2 molecule of 414 amino acids with Glu at
XX	CC	its NH2-terminus. See ARI1128 for PAI-2 with a synthetic signal
XX	CC	sequence. (Updated on 09-JAN-2003 to add missing OS field.)
XX	XX	Sequence 1340 BP; 374 A; 307 C; 318 G; 341 T; 0 U; 0 Other;
XX	XX	Query Match 65.1%; Score 1242.8; DB 2; Length 1340;
XX	XX	Best Local Similarity 99.4%; Pred. No. 1.5e-247;
XX	XX	Matches 1247; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY	56	GAGGATCTTTGTGTGGCAACACACTCTTTCGCCCTCAATTTATTTCAGCATCTGGCAAAA 125
DB	87	GAGGATCTTTGTGTGGCAACACACTCTTTCGCCCTCAATTTATTTCAGCATCTGGCAAAA 146
QY	126	GCAAGGCCCAACCCAGAACCTTCTCTCCCATGGAGCATCTGTCACCATGGCCATG 185
DB	147	GCAAGGCCCAACCCAGAACCTTCTCTCTCCCATGGAGCATCTGTCACCATGGCCATG 206
QY	186	GTCTACATGGCTCCAGGGGAGCAGCAGGAAGACCATGGCCCAAGGTGCTTCAGTTTAAAT 245
DB	207	GTCTACATGGCTCCAGGGGAGCAGCAGGAAGACCATGGCCCAAGGTGCTTCAGTTTAAAT 266
QY	246	GAAAGTGGAGCCAAATGCAAGTTACCCCATGACTCCAGAGAACTTTACAGCTGTGGGTTT 305
DB	267	GAAAGTGGAGCCAAATGCAAGTTACCCCATGACTCCAGAGAACTTTACAGCTGTGGGTTT 326
QY	306	ATGACGAGATCCAGAGGGTAGTTATCTCTGATGCGATTTTCAGGCAAAAGCTGCAGAT 365
DB	327	ATGACGAGATCCAGAGGGTAGTTATCTCTGATGCGATTTTCAGGCAAAAGCTGCAGAT 386
QY	366	AAATCCATTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTAT 425
DB	387	AAATCCATTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTAT 446
QY	426	TTACTGAAAGTGTCAATAAGCTGTTTGGTAGAAGCTCTGGAGCTTCCGGGAAGATAT 485
DB	447	TTACTGAAAGTGTCAATAAGCTGTTTGGTAGAAGCTCTGGAGCTTCCGGGAAGATAT 506
QY	486	ATTTCGACTCTGTCAAGAAATATTACTCTCTCAGAACCCCGAGGAGTAGACTTCTTAGAATGT 545
DB	507	ATTTCGACTCTGTCAAGAAATATTACTCTCTCAGAACCCCGAGGAGTAGACTTCTTAGAATGT 566
QY	546	GCAGAGAACTAGAAAAAGATTAATTTCTCTGGGTCAAGACTCAAAACCAAGGCAAAATC 605
DB	567	GCAGAGAACTAGAAAAAGATTAATTTCTCTGGGTCAAGACTCAAAACCAAGGCAAAATC 626

```
2Y 606 CCAAACTTGTACCTGAAGGTTCTGTAGATGGGATACCAAGATGGTCTCTGTGAATGCT 665
Db 627 CCAAACTTGTACCTGAAGGTTCTGTAGATGGGATACCAAGATGGTCTCTGTGAATGCT 686
2Y 666 GTCTACTTCAAGGAAAGTGGAAAACCTCAATTGAGAAGAACTAAATGGGCTTTATCCT 725
Db 687 GTCTACTTCAAGGAAAGTGGAAAACCTCAATTGAGAAGAACTAAATGGGCTTTATCCT 746
2Y 726 TTCGTGTAACTCGGCTCAGCGCACCTGTACAGATGATGTACTTTCGTGAAAAGCTA 785
Db 747 TTCGTGTAACTCGGCTCAGCGCACCTGTACAGATGATGTACTTTCGTGAAAAGCTA 806
QY 786 AACATTGATACATAGAAAGACCTAAAGGCTCAGATTCAGAACTCCATATGCTGGAGAT 845
Db 807 AACATTGATACATAGAAAGACCTAAAGGCTCAGATTCAGAACTCCATATGCTGGAGAT 866
QY 846 GTTAGCATGTTCTTGTGCTTCAGATGAATGCGGATGTGCTCACTGGCTTGGAGCTG 905
Db 867 GTTAGCATGTTCTTGTGCTTCAGATGAATGCGGATGTGCTCACTGGCTTGGAGCTG 926
QY 906 CTGGAAGTGAATTAACCTATGACAACTCAACAAGTGGACCAAGCAAAATGGCT 965
Db 927 CTGGAAGTGAATTAACCTATGACAACTCAACAAGTGGACCAAGCAAAATGGCT 986
QY 966 GAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAGAGCATTAAGAACTCAGA 1025
Db 987 GAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAGAGCATTAAGAACTCAGA 1046
QY 1026 TCCATTCTGAAAGCATGGCGATGGAGGAGCGCTTCAACAAGGAGCGGGCAATTTCTCA 1085
Db 1047 TCCATTCTGAGAGCATGGCGATGGAGGAGCGCTTCAACAAGGAGCGGGCAATTTCTCA 1106
QY 1086 GGGATGTGGAGAGGATGACCTGTTCTTCTTGAAGTGTTCACCAAGCATGCTGGAT 1145
Db 1107 GGGATGTGGAGAGGATGACCTGTTCTTCTTGAAGTGTTCACCAAGCATGCTGGAT 1166
QY 1146 GTGAATGAGAGGCGCACTGAAGCAGCGCTGGCAGAGAGGTGTATGACAGGAGAACT 1205
Db 1167 GTGAATGAGAGGCGCACTGAAGCAGCGCTGGCAGAGAGGTGTATGACAGGAGAACT 1226
QY 1206 GGACATGGAGGCGCCACAGTTGTGGCAGATCATCCGTTCTTTCTTATATGATCAAG 1265
Db 1227 GGACATGGAGGCGCCACAGTTGTGGCAGATCATCCGTTCTTTCTTATATGATCAAG 1286
QY 1266 ATAACCAAGTGCATTTTATTTTTCGAGATTTTGTCCACCCCTAAACTAAGCG 1319
Db 1287 ATAACCAAGTGCATTTTATTTTTCGAGATTTTTCCTCACCCCTAAACTAAGCG 1340
```

RESULT 14

AAQ21119 standard; cDNA; 1290 BP.

XX AAQ21119;

XX AC

XX 25-MAR-2003 (revised)

DT 26-MAY-1992 (first entry)

XX PAI-2 in pDBP3.

XX Plasminogen activator inhibitor; PAI; protease; uPA; receptor;

KW inhibition; urokinase; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FH 25..1272

CDS /*tag= a

FT /label= PAI-2

XX GB2246779-A.

XX

XX

12-FEB-1992.

XX 03-AUG-1990; 90GB-00017083.

XX 03-AUG-1990; 90GB-00017083.

XX (DELZ) DELTA BIOTECH LTD.

XX Ballance DJ, Courtney MG;

XX WPI; 1992-051155/07.

XX P-PSDB; AAR20797.

XX Antitumour molecules for treatment of neoplasms - comprises first region

XX for binding to uPA receptor and second region for uPA inhibition.

XX Disclosure; Fig 2; 57pp; English.

XX A PAI-2 positive clone was isolated after screening a library (see
XX AAQ21117-8). A BglII site was created at the 5' end of the gene to create
XX a mutation in the third position of the second codon. A AlfrI site was
XX created at the 3' end to create mutations in the third position of the
XX last codon and in the first base of the stop codon. A clone derived from
XX this procedure was designated pDBP2. Two linkers were ligated with the
XX BglII-AfrII PAI-2 gene fragment from pDBP2 into HindIII + BamHI digested
XX m13mp19 to form pDBP3. See also AAQ21117-19 and AAQ21121-25. (Updated on
XX 25-MAR-2003 to correct PA field.)

SQ Sequence 1290 BP; 377 A; 289 C; 305 G; 319 T; 0 U; 0 Other;

Query Match 65.1%; Score 1242; DB 2; Length 1290;

Best Local Similarity 99.2%; Pred. No. 2.2e-247;

Matches 1248; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 58 AAACAATGGAGATCTTGTGGCAACACACACTTTGGCCCTCAATTTATTCAAGCATC 117

Db 20 AAAAATGGAAGATCTTGTGGCAACACACTTTGGCCCTCAATTTATTCAAGCATC 79

QY 118 TGGCAAAAGCAAGCCCAACCCAGAAACCTCTTCTCTCCCATGAGCATCTCGTCCACA 177

Db 80 TGGCAAAAGCAAGCCCAACCCAGAAACCTCTTCTCTCCCATGAGCATCTCGTCCACA 139

QY 178 TGGCCATGTCTACATGGGCTCCAGGGGAGCAGCCAGAACCCAGATGCCAAGGTGCTTC 237

Db 140 TGGCCATGTCTACATGGGCTCCAGGGGAGCAGCCAGAACCCAGATGCCAAGGTGCTTC 199

QY 238 AGTTTAATGAAGTGGAGCCCAATGCGATTTACCCCATGACTCCAGAGNACTTTACCAGCT 297

Db 200 AGTTTAATGAAGTGGAGCCCAATGCGATTTACCCCATGACTCCAGAGNACTTTACCAGCT 259

QY 298 GTGGGTTTCATGAGCAGATCCAGAAGGGTAGTTATCTCTGATGCGATTTTGAGGCACAAG 357

Db 260 GTGGGTTTCATGAGCAGATCCAGAAGGGTAGTTATCTCTGATGCGATTTTGAGGCACAAG 319

QY 358 CTGCAGATAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATCAATCCACAG 417

Db 320 CTGCAGATAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATCAATCCACAG 379

QY 418 GGAATTAATTTACTGGAAAGTGTCAATAAGCTGTTGTTGAGAAAGTCTCGAGCTTCCGGG 477

Db 380 GGAATTAATTTACTGGAAAGTGTCAATAAGCTGTTGTTGAGAAAGTCTCGAGCTTCCGGG 439

QY 478 AAGAATATATTCGACTCTGTGAGAAATATTACTCTCAGAACCCAGCAGTAGACTTCC 537

Db 440 AAGAATATATTCGACTCTGTGAGAAATATTACTCTCAGAACCCAGCAGTAGACTTCC 499

QY 538 TAGAATGTGAGAAAGCTAGAAAAGATTAAATTTCTGGTCAAGACTCAAAACCAAG 597

Db 500 TAGAATGTGAGAAAGCTAGAAAAGATTAAATTTCTGGTCAAGACTCAAAACCAAG 559

QY 598 GCAAAATCCCAACTTGTATTACCTGAAGGTTCTGTAGATGGGATACCAAGGATGGTCTCGG 657

Db 560 GCAAAATCCCAACTTGTATTACCTGAAGGTTCTGTAGATGGGATACCAAGGATGGTCTCGG 619

658	QY	TGAATGCTGCTCTACTTTCAAAGGAAAGTGAAAACTCCATTTGAGAAAGAAACTAAATGGGC	717
620	Db	TGAATGCTGCTCTACTTTCAAAGGAAAGTGAAAACTCCATTTGAGAAAGAAACTAAATGGGC	679
718	QY	TTTATCCTTTCCGTGTAAACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTTCGGTG	777
680	Db	TTTATCCTTTCCGTGTAAACTCGGCTCAGCGCACACCTGTACAGATGATGTACTTTCGGTG	739
778	QY	AAAAGCTAAACATTGATACATAGAAGACTAAAGGCTCAGATTTAGAACTCCCATATG	837
740	Db	AAAAGCTAAACATTGATACATAGAAGACTAAAGGCTCAGATTTAGAACTCCCATATG	799
838	QY	CTGGAGATGTTAGCATGTTCTTGTTCCTTCCAGATGAAATTGGCGATGTGTCCACTGGCT	897
800	Db	CTGGAGATGTTAGCATGTTCTTGTTCCTTCCAGATGAAATTGGCGATGTGTCCACTGGCT	859
898	QY	TGGAGTGTCTGGAAGTGAATAACCTATGACAAACTCAACAGTGGACAGCAAAAGACA	957
860	Db	TGGAGTGTCTGGAAGTGAATAACCTATGACAAACTCAACAGTGGACAGCAAAAGACA	919
958	QY	AAATGGCTCAAGATGAAGTTGAGGTATACATACCCAGTTTCAAATTTAGAAGACATTTATG	1017
920	Db	AAATGGCTCAAGATGAAGTTGAGGTATACATACCCAGTTTCAAATTTAGAAGACATTTATG	979
1018	QY	AACTCAGATCCATTTCTGAAAAGCATGGGATGAGGACGCTTCAACAAGGNCGGGCCA	1077
980	Db	AACTCAGATCCATTTCTGAAAAGCATGGGATGAGGACGCTTCAACAAGGNCGGGCCA	1039
1078	QY	ATTTCCTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTTCTGAAGTGTTCACCAAGCCA	1137
1040	Db	ATTTCCTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTTCTGAAGTGTTCACCAAGCCA	1099
1138	QY	TGTTGATGTGAATGAGGAGGGCACTGAAGCAGCGCTGGCACAGGAGGTGTTATGACAG	1197
1100	Db	TGTTGATGTGAATGAGGAGGGCACTGAAGCAGCGCTGGCACAGGAGGTGTTATGACAG	1159
1198	QY	GGAAACTGGAATGAGAGGCCCAAGTTTGTGGCAGATCATCGTTTCTTTTCTTTATTA	1257
1160	Db	GGAAACTGGAATGAGAGGCCCAAGTTTGTGGCAGATCATCGTTTCTTTTCTTTATTA	1219
1258	QY	TGCATGAAGATAACCAAGTGCATTTATTTTTCGGCAGATTTTGTCTCACCTCAAAACTA	1315
1220	Db	TGCATGAAGATAACCAAGTGCATTTATTTTTCGGCAGATTTTGTCTCACCTCAAGTCGA	1277

RESULT 15

ABZ83797
ID ABZ83797 standard; cDNA: 1024 BP.

XX
AC ABZ83797:

XX
DT 14-MAY-2003 (first entry)

XX Toxicologically relevant human nucleotide sequence #956.

XX
KW
XX

XX
QS Homo sapiens.

XX PN WO2003016500-A2.

XX
PD 27-FEB-2003.

XX
PF 16-AUG-2002: 2002WO-US026514.

XX
PR 16-AUG-2001; 2001US-0313080P.

XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.

XX PI Neft RE. Dunn RT. Adkins K. Pickett GG, Kier LD, Schmeiser K;

PI Alen P;
XX

WPI; 2003-268322/26.

Determining a toxicological response to an agent, useful for screening of drugs, comprises comparing the expression profile of one or more human toxic response genes to a reference gene expression profile indicative of toxicity.

Claim 1; Page 298; 455pp; English.

The present invention describes a method (M1) for determining a toxicological response to an agent, which comprises comparing the expression profile of one or more human toxic response genes to a reference gene expression profile indicative of toxicity, and so determining the presence of a toxic response to the agent. Also described: (1) an array comprising one or more polynucleotides selected from the genes corresponding to the partial sequences given in ABZ82842 to ABZ94754, or their fragments of at least 20 nucleotides, or homologues; and (2) determining if a gene putatively identified to be a toxic response gene plays a role on toxic response pathways by determining the expression profile of the gene after exposure of cells or a human subject to a known toxic pharmaceutical or industrial agent, comprising: (a) exposing cells to an agent or isolating cells from a human subject who was exposed to an agent; (b) obtaining the test gene expression profile for a putatively identified toxic response gene after exposure to a known toxic pharmaceutical or industrial agent; and (c) comparing the test profile to the expression profile of a gene with a similar function or comparing the test profile to the expression profile of that gene after exposure to other known toxic compounds. The methods are useful for predicting and determining toxicological responses on a cellular, organ or system level. The arrays comprising the human genes are useful for toxicological screening of drugs, pharmaceutical compounds and chemicals.

Sequence 1024 BP: 313 A; 233 C; 234 G; 244 T; 0 U; 0 Other;

Query Match 53.1%; Score 1014; DB 7; Length 1024;

```

Query Match 55.1%; Score 1017; Length 1000
Best Local Similarity 99.5%; Pred. No. 3.3e-200;
Matches 1017; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

34 AATAACCAGAGAACAAACCAGATTGAAACAAATGGAGGATCTTTGTGTGGCAACACACTCT 93

3 ATCCCCAGAGAACAAACCAGATTGAAACAATGGAGGATCTTTGTGTGGCAACACACTCT 62

94 TTCCCCCTCAATTATTTCAAGCATCTGGCAAAAGCAAGCCCAAGCCCTCTTCTCTT 153

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

QY
I54 CCCCATTGGAGCAATCCTCGTCCACCATGAGGCCAATGGTCATACTATGGGCCTCCACCGCCCGAACACGCTTT

Db 123 CCCCATGGAGCATCTCGTCCACCATGGCCATGGTCTACATGGGCTCCAGGGCAGCACCG 182

214 AAGACCATGGCCCAAGGTGCTTCAGTTTAAATGAAGTGGAGCCAATGCAGTTACCCCCA .273 QY

db
183 AAGACCAGATGGCCCAAGGTGCTTCAGTTAATGAAGTGGAGCCAAATGCAGTTACCCCCA 242

274 TGA CTC CAG AGA A C T T T A C C A G C T G T G G G T T C A T G C A G C A G A T C C A G A A G G T A G T A T C 333

302

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

[illegible]

Db 303 CTGATGCGATTTTGCAGGCACAAGCTGCAGAT/AAAATCCATTCATCCCTCCGCTCTCTCA 308

394 GCTCTGCAATCAATGCATCCACAGGAATTATTACTGGAAAGTGTCAATAAGCTGTTTG 45

Db 363 GCTCTGCAATCAATGCATCCACAGGGGATTATTTACTGGAAAGTGTCATAAGCTGTTG 422

454 GTGAGAAGTCTCGGAGCTTCGGGAAGAAATATATTCGACTCTGTCTCAGAAATATTACTCCT 513

423 CTGCGAGCTCGCGAGCTTCCGGGAGAAATATATTCGACTCTGTTCAGAAATATTACTCCT 482


```

y 574 CTTGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACCTGTTACCTCAAGGTTCTGTAG 633
b 543 CTTGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACCTGTTACCTCAAGGTTCTGTAG 602
y 634 ATGGGATACAGGATGGTCTGTTGAATGCTGCTACTTCAAAGGAAAGTGGAAACTC 693
b 603 ATGGGATACAGGATGGTCTGTTGAATGCTGCTACTTCAAAGGAAAGTGGAAACTC 662
y 694 CATTTGAGAAAGAACTAAATGGGCTTTATCCCTTCCGTTGTAACCTCGGCTCAGCCACAC 753
b 663 CATTTGAGAAAGAACTAAATGGGCTTTATCCCTTCCGTTGTAACCTCGGCTCAGCCACAC 722
y 754 CTGTACAGATGATGTACTTGGTGAAGCTAAACATTGGATACATAGAAGACCTAAAGG 813
b 723 CTGTACAGATGATGTACTTGGTGAAGCTAAACATTGGATACATAGAAGACCTAAAGG 782
y 814 CTCAGATTTAGAACTCCCATATGCTGGAGATGTAGCATGTTCTTGTGCTTCCAGATG 873
b 783 CTCAGATTTAGAACTCCCATATGCTGGAGATGTAGCATGTTCTTGTGCTTCCAGATG 842
y 874 AAATTGCCGATGTGCCACTGGCTTGGAGCTGCTGGAAGTGAATAACCTATGACAAAC 933
b 843 AAATTGCCGATGTGCCACTGGCTTGGAGCTGCTGGAAGTGAATAACCTATGACAAAC 902
y 934 TCAACAAGTGGACCAAGCAAAATGGCTGAAGATGAAGTTGAGGTATACATACCCC 993
b 903 TCAACAAGTGGACCAAGCAAAATGGCTGAAGATGAAGTTGAGGTATACATACCCC 962
y 994 AGTTCAAATTAGAAGAGCATTTATGAACCTCAGATCCATTCTGAAAAGCATGGGCATGGAGG 1053
b 963 AGTTCAAATTAGAAGAGCATTTATGAACCTCAGATCCATTCTGAGAAGCATGGGCATGGAGG 1022
y 1054 AC 1055
b 1023 AC 1024

```

Search completed: March 10, 2004, 10:59:51
Job time : 471.37 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:19:42 ; Search time 87.1303 Seconds
(without alignments)
12152.440 Million cell updates/sec

Title: US-10-084-817-27

Perfect score: 1908

Sequence: 1 gagagcatgcccgtcaga.....aacaaaaataaaaaagg 1908

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A-COMB.seq*
- 2: /cgn2_6/ptodata/2/ina/5B-COMB.seq*
- 3: /cgn2_6/ptodata/2/ina/6A-COMB.seq*
- 4: /cgn2_6/ptodata/2/ina/6B-COMB.seq*
- 5: /cgn2_6/ptodata/2/ina/PTUS-COMB.seq*
- 6: /cgn2_6/ptodata/2/ina/backfile1.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1876.6	98.4	1884	4	US-09-023-655-1188
2	1847.6	96.8	1854	4	US-09-023-655-1189
3	1847.6	96.8	2409	1	US-07-911-531-18
4	1847.6	96.8	2409	1	US-07-933-636A-18
5	1802	84.0	1810	1	US-07-968-286B-1
6	1402.6	73.5	1512	1	US-07-768-286B-3
7	1342.6	70.4	1482	1	US-07-768-286B-5
8	1249.4	65.3	1328	1	US-07-679-052A-14
9	1246	65.3	1340	1	US-07-679-052A-16
10	279	14.6	1519	4	US-10-024-427-3
11	274	14.4	1278	4	US-10-024-427-1
12	268.4	14.1	1316	1	US-08-315-831A-12
13	268.4	14.1	1316	1	US-08-315-831A-12
14	268.4	14.1	1316	5	PCT-US95-12509-12
15	268.4	14.1	1430	4	US-09-016-434-710
16	266.8	14.0	1316	1	US-07-755-461A-12
17	265.8	13.9	1152	1	US-08-315-831A-13
18	265.8	13.9	1152	1	US-08-662-318-13
19	265.8	13.9	1152	5	PCT-US95-12509-13
20	252.4	13.2	1385	4	US-09-016-434-1075
21	252.4	13.2	1385	4	US-09-023-655-892
22	252.4	13.2	1393	1	US-08-464-148-3
23	252.4	13.2	1393	1	US-08-385-500-3
24	252.4	13.2	1393	1	US-08-846-784-3
25	251.6	13.2	1366	3	US-09-200-965-1
26	249.2	13.1	1361	4	US-09-016-434-1284
27	237.4	12.4	241	4	US-09-389-681-338

C

C	28	237.4	12.4	241	4	US-09-620-405B-338	Sequence 338, App
C	29	237.4	12.4	241	4	US-09-433-826B-338	Sequence 338, App
C	30	237.4	12.4	241	4	US-09-604-287A-338	Sequence 338, App
C	31	237.4	12.4	241	4	US-09-834-755-338	Sequence 338, App
	32	234.8	12.3	1308	4	US-09-123-912-109	Sequence 109, App
	33	234.8	12.3	1308	4	US-09-643-597-109	Sequence 109, App
	34	234.8	12.3	1308	4	US-09-480-884A-109	Sequence 109, App
	35	234.8	12.3	1308	4	US-09-542-615A-109	Sequence 109, App
	36	234.8	12.3	1308	4	US-09-606-421B-109	Sequence 109, App
	37	234.8	12.3	1308	4	US-09-221-107-109	Sequence 111, App
	38	234.8	12.3	1419	4	US-09-123-912-111	Sequence 111, App
	39	234.8	12.3	1419	4	US-09-643-597-111	Sequence 111, App
	40	234.8	12.3	1419	4	US-09-480-884A-111	Sequence 111, App
	41	234.8	12.3	1419	4	US-09-542-615A-111	Sequence 111, App
	42	234.8	12.3	1419	4	US-09-606-421B-111	Sequence 111, App
	43	234.8	12.3	1419	4	US-09-221-107-111	Sequence 111, App
	44	229.8	12.0	1245	4	US-09-266-910-1	Sequence 1, Appli
	45	229.4	12.0	1711	1	US-08-568-147B-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-023-655-1188
; Sequence 1188, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1188:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1884 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENEBANK
; CLONE: g189544
US-09-023-655-1188

Query Match

98.4%; Score 1876.6; DB 4; Length 1884;

[illegible]

1022	ATCTCAGAGCATGGGCATGGAGGACGCCTTCAACAGGGACGGGCCAATTTCTCAGG	1081
1089	ATGTCGGAGAGGAATGACCTGTTCTTCTTCGAAAGTGTTCCACCAAGCAGTGGTGAATGG	1148
1082	ATGTCGGAGAGGAATGACCTGTTCTTCTTCGAAAGTGTTCCACCAAGCAGTGGTGAATGG	1141
1149	AATGAGGAGGGCACTGAAGCAGCGCTGGCACAAGGAGGTGTTATGACAGGGAGAACTGGA	1208
1142	AATGAGGAGGGCACTGAAGCAGCGCTGGCACAAGGAGGTGTTATGACAGGGAGAACTGGA	1201
1209	CATGGAGGCCACAGTTTGTGGCAGATCATCCGTTTCTTTTCTTATTAATGCAATAGATA	1268
1202	CATGGAGGCCACAGTTTGTGGCAGATCATCCGTTTCTTTTCTTATTAATGCAATAGATA	1261
1269	ACCAAGTGCATTTATTTTTTGGCAGATTTGTCTCACCCCTAAACCTAAGCGTCTGCTTC	1328
1262	ACCAAGTGCATTTATTTTTTGGCAGATTTGTCTCACCCCTAAACCTAAGCGTCTGCTTC	1321
1329	TGCAAAAGATTTTGTAGATGAGCTGTGCGCTCAGAAATGCTATTTCAAATGCCCAAA	1388
1322	TGCAAAAGATTTTGTAGATGAGCTGTGCGCTCAGAAATGCTATTTCAAATGCCCAAA	1381
1389	ATTTAGAGATGTTTTCTACATATTTCTGCTCTTCTGAAACAACCTCTGTACCCATAAT	1448
1382	ATTTAGAGATGTTTTCTACATATTTCTGCTCTTCTGAAACAACCTCTGTACCCATAAT	1441
1449	AAAAACACAGAAATTAATAGACAAATGCTCTATTATACATGACAAACCTATTAAATCATTT	1508
1442	AAAAACACAGAAATTAATAGACAAATGCTCTATTATACATGACAAACCTATTAAATCATTT	1501
1509	GGTCTTCTAAAATGGGATCATGCCCATTTAGATTTTCTTCTTACTATACAGTTATTTTATA	1568
1502	GGTCTTCTAAAATGGGATCATGCCCATTTAGATTTTCTTCTTACTATACAGTTATTTTATA	1561
1569	ACATTAACTTTTACTTTGTTTATTTATTTTATATAAATGGTCAGATTTTAAATATTGTC	1628
1562	ACATTAACTTTTACTTTGTTTATTTATTTTATATAAATGGTCAGATTTTAAATATTGTC	1621
1629	TCAGTCCTATTTAAATGTAGCTAATAAAGTTATAGAAGCAGATGATCTGTTAAATTTCCCTA	1688
1622	TCAGTCCTATTTAAATGTAGCTAATAAAGTTATAGAAGCAGATGATCTGTTAAATTTCCCTA	1681
1689	TCTAAATAATGTCCTTTAATGTTTCTCATATGAAGAATAAGTAGGTATCCCTCCATGCC	1748
1682	TCTAAATAATGTCCTTTAATGTTTCTCATATGAAGAATAAGTAGGTATCCCTCCATGCC	1741
1749	TTCTGTAATAAATATCTGGAAAAAACAATTAAACAATAGGCAATATATGTTATGTGCATT	1808
1742	TTCTTAATAATAATCTGGAAAAAACAATTAAACAATAGGCAATATATGTTATGTGCATT	1801
1809	TCTAGAATAACATAACACATATATGTCGTGATCTTATATTCAAATGCAAGTATATAAT	1868
1802	TCTAGAATAACATAACACATATATGTCGTGATCTTATATTCAAATGCAAGTATATAAT	1861
1869	AAATAAACCTGCTTCCAAAACAC	1891
1862	AAATAAACCTGCTTCCAAAACAC	1884

RESULT 2
US-09-023-655-1189
; Sequence 1189, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE

CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1189:
SEQUENCE CHARACTERISTICS:
LENGTH: 1854 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g189546
(S-09-023-655-1189)

Query Match 96.8%; Score 1847.6; DB 4; Length 1854;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

15	GTCTGCGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCC	74
1	GTCTGCGCAACACTCTTTGGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCC	60
75	TGCTGCGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCC	134
61	TGCTGCGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCC	120
135	ACCAGAACCTTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGTCTCATG	194
121	ACCAGAACCTTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGTCTCATG	180
195	GGCTCCAGGGGCGAGCACCGAAGACAGATGGCAAGTGCTTCAGTTTAAAGTGGGA	254
181	GGCTCCAGGGGCGAGCACCGAAGACAGATGGCAAGTGCTTCAGTTTAAAGTGGGA	240
255	GCCAAATGAGTTACCCCATGACTCCAGAGAACTTTACAGCTGTGGTTCATGAGCAG	314
241	GCCAAATGAGTTACCCCATGACTCCAGAGAACTTTACAGCTGTGGTTCATGAGCAG	300
315	ATCCAGAGGGGTAGTTATCTCTGATGGATTTTGCAGGCAAGCTCGAGTAAATCCAT	374
301	ATCCAGAGGGGTAGTTATCTCTGATGGATTTTGCAGGCAAGCTCGAGTAAATCCAT	360
375	TCATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGAATTTATTTACTGGAA	434
361	TCATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGAATTTATTTACTGGAA	420
435	AGTGTCAATAAGCTGTTTGGTGAGAGTCTGGGAGCTTCCGGGAAGATATATTCGACTC	494
421	AGTGTCAATAAGCTGTTTGGTGAGAGTCTGGGAGCTTCCGGGAAGATATATTCGACTC	480
495	TGTCAGAAATATTACTCTCTCAGAACCCAGGCGAGTAGACTTCTCTAGAAATGTGAGAGAA	554

Db	481	TGTCAGAAATATTACTCTCTCAGAACCCAGGCGAGTAGACTTCTCTAGAAATGTGAGAGAA	540
Qy	555	GCTAGAAAAAGATTATTTCTGGTCAAGACTCAAAACAAAGGCAAAATCCCAAACTTG	614
Db	541	GCTAGAAAAAGATTATTTCTGGTCAAGACTCAAAACAAAGGCAAAATCCCAAACTTG	600
Qy	615	TTACTGAAGGTTCTGTAGATGGGATACAGAGATGGTCTGTGTGAATGTGTCTACTTC	674
Db	601	TTACTGAAGGTTCTGTAGATGGGATACAGAGATGGTCTGTGTGAATGTGTCTACTTC	660
Qy	675	AAAGGAAAGTGGAAAACTCCATTTGAGAGAACTAAATGGGCTTTATCTTTCCGTGTA	734
Db	661	AAAGGAAAGTGGAAAACTCCATTTGAGAGAACTAAATGGGCTTTATCTTTCCGTGTA	720
Qy	735	AACCTGGCTCAGGCGACACCTGTGTA CAGATGATGTTACTTGGTGAAAAAGCTAAACATTGGA	794
Db	721	AACCTGGCTCAGGCGACACCTGTGTA CAGATGATGTTACTTGGTGAAAAAGCTAAACATTGGA	780
Qy	795	TACATAGAAAGACCTAAAGGCTCAGATTCTAGAACTCCCATATGCTGGAGATGTTAGCATG	854
Db	781	TACATAGAAAGACCTAAAGGCTCAGATTCTAGAACTCCCATATGCTGGAGATGTTAGCATG	840
Qy	855	TTCTTTGTTGTTCCAGATGAATTCGCGATGTGCCACTGGCTTGGAGCTGCTGGAAGT	914
Db	841	TTCTTTGTTGTTCCAGATGAATTCGCGATGTGCCACTGGCTTGGAGCTGCTGGAAGT	900
Qy	915	GAATAACCTATGACAAACTCAACAGTGGACAGCAAAAGACAAAATGGCTGAAGATGAA	974
Db	901	GAATAACCTATGACAAACTCAACAGTGGACAGCAAAAGACAAAATGGCTGAAGATGAA	960
Qy	975	GTTGAGGTATACATACCCAGTTCAANTTAGAGAGCATTTAGACTCAGATCCATCTG	1034
Db	961	GTTGAGGTATACATACCCAGTTCAANTTAGAGAGCATTTAGACTCAGATCCATCTG	1020
Qy	1035	AAAAAGCATGGGATGGAGACGCCCTTCAACAAAGGAGCGGCCCAATTTCTCAGGGATGTC	1094
Db	1021	AGAAAGCATGGGATGGAGACGCCCTTCAACAAAGGAGCGGCCCAATTTCTCAGGGATGTC	1080
Qy	1095	GAGAGGAATGACCTTTCTTCTCAAGTGTCCACCAAGCCCATGGTGGATGTAAGTAG	1154
Db	1081	GAGAGGAATGACCTTTCTTCTCAAGTGTCCACCAAGCCCATGGTGGATGTAAGTAG	1140
Qy	1155	GAGGCACTGAAGCAGCGCTGGCA CAGGAGTGTATTGACAGGGAGAACTGGACATGGA	1214
Db	1141	GAGGCACTGAAGCAGCGCTGGCA CAGGAGTGTATTGACAGGGAGAACTGGACATGGA	1200
Qy	1215	GGCCCAAGTTTGTGGCAGATCATCGTTTCTTTCTTTTCTTTTATGATAGATAACCAAG	1274
Db	1201	GGCCCAAGTTTGTGGCAGATCATCGTTTCTTTCTTTTCTTTTATGATAGATAACCAAG	1260
Qy	1275	TGCATTTTATTTTTCGGCAGATTTTGTCTCA CCGTAAACCTAAGCGTCTGCTTCTGCAAA	1334
Db	1261	TGCATTTTATTTTTCGGCAGATTTTGTCTCA CCGTAAACCTAAGCGTCTGCTTCTGCAAA	1320
Qy	1335	AGATTTTGTAGATGAGCTGTGCTCAGATTTGCTATTTTCAAAATGCCAAAAATTTAG	1394
Db	1321	AGATTTTGTAGATGAGCTGTGCTCAGATTTGCTATTTTCAAAATGCCAAAAATTTAG	1380
Qy	1395	AGATGTTTCTTACATATTTCTCTCTTGAACAACTCTGTGTA CCGTAAATFAAAAC	1454
Db	1381	AGATGTTTCTTACATATTTCTCTCTTGAACAACTCTGTGTA CCGTAAATFAAAAC	1440
Qy	1455	ACAGAAATATTAGACAATTTGCTATTATAA CAGTGAACACCCCTATTATCATTTGGTCTT	1514
Db	1441	ACAGAAATATTAGACAATTTGCTATTATAA CAGTGAACACCCCTATTATCATTTGGTCTT	1500
Qy	1515	CTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCATCAGTTATTTTATTAACATTA	1574
Db	1501	CTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCATCAGTTATTTTATTAACATTA	1560
Qy	1575	ACTTTTACTTGTATTATTTATTTATTAATGAGTGGATGAGTTTATTAATTTTGTCTCACTG	1634
Db	1561	ACTTTTACTTGTATTATTTATTTATTAATGAGTGGATGAGTTTATTAATTTTGTCTCACTG	1620

1635 CCTATTATAGTACCTAATAAGATTATAGAGCAGATGATCTGTTAAATTCCTATCTAAT 1694
1621 CCTATTATAGTACCTAATAAGATTATAGAGCAGATGATCTGTTAAATTCCTATCTAAT 1690
1695 AAATGCCCTTAAATGTTCTCATAATGAAGAATAAGTAGGTATCCCTCCATGCCCTTCTGT 1754
1681 AAATGCCCTTAAATGTTCTCATAATGAAGAATAAGTAGGTATCCCTCCATGCCCTTCTGT 1740
1755 AATAAATATCTGGAAAAACATTAACAAATAGCAATATATGTTATGTCATTTCTAGA 1814
1741 AATAAATATCTGGAAAAACATTAACAAATAGCAATATATGTTATGTCATTTCTAGA 1800
1815 AATACATAACACATATATATGTCATTTCTATATTAATCAATGCAAGTATATAAT 1868
1801 AATACATAACACATATATATGTCATTTCTATATTAATCAATGCAAGTATATAAT 1854

RESULT 3

US-07-911-531-18

; Sequence 18, Application US/07911531

; Patent No. 5422090

; GENERAL INFORMATION:

; APPLICANT: STEPHENS, Ross

; APPLICANT: GOLDBER, Jeffrey

; APPLICANT: ANTALIS, Toni M.

; APPLICANT: BARNES, Thomas M.

; APPLICANT: CLARK, Michelle A.

; APPLICANT: DEVINE, Peter L.

; APPLICANT: GOSS, Neil H.

; APPLICANT: LEHRBACH, Philip R.

; TITLE OF INVENTION: MINACTIVIN

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500

; CITY: Alexandria

; STATE: Virginia

; COUNTRY: USA

; ZIP: 22313-0299

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/911,531

; FILING DATE: 19920715

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 16786/165 CHAC

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)836-9300

; TELEFAX: (703)836-4109

; TELEX: 899149

; INFORMATION FOR SEQ ID NO: 18:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2409 base pairs

; TYPE: NUCLEIC ACID

; STRANDEDNESS: double

; TOPOLOGY: unknown

; MOLECULE TYPE: cDNA to genomic RNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 49..1296

; US-07-911-531-18

Query Match

96.8%; Score 1847.6; DB 1; Length 2409;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 15 GTCAGACAGCAACTCAGAGATAAACCCAGAGAACCAACAGATTGAAACAATGGAGATCTT 74
DB 1 GTCAGACAGCAACTCAGAGATAAACCCAGAGAACCAACAGATTGAAACAATGGAGATCTT 60
QY 75 TGTGTGGCAAAACACACTCTTTTGGCCCTCAATTTTCAAGCATCTGGCAAAAGCAAGCCCC 134
DB 61 TGTGTGGCAAAACACACTCTTTTGGCCCTCAATTTTCAAGCATCTGGCAAAAGCAAGCCCC 120
QY 135 ACCAGAACCTTCTTCTCCCATCGAGCATCTGTCACCATCGGCATGGTCTACATG 194
DB 121 ACCAGAACCTTCTTCTCCCATCGAGCATCTGTCACCATCGGCATGGTCTACATG 180
QY 195 GGTCCAGGGGCGAGCACCGAAGACCAAGTGGCCAAAGTGTCTTCAATTTAATGAAGTGGGA 254
DB 181 GGTCCAGGGGCGAGCACCGAAGACCAAGTGGCCAAAGTGTCTTCAATTTAATGAAGTGGGA 240
QY 255 GCCAATGCAATACCCCATGACTCCAGAGAACTTTACCAAGCTGTGGTTCATGCAGCAG 314
DB 241 GCCAATGCAATACCCCATGACTCCAGAGAACTTTACCAAGCTGTGGTTCATGCAGCAG 300
QY 315 ATCCAGAGGGTAGTTATCCTGATGCGATTTTCAGGCACAAGCTGCAGATAAAATCCAT 374
DB 301 ATCCAGAGGGTAGTTATCCTGATGCGATTTTCAGGCACAAGCTGCAGATAAAATCCAT 360
QY 375 TCATCTCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCAGGGAATTTACTGGAA 434
DB 361 TCATCTCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCAGGGAATTTACTGGAA 420
QY 435 AGTGCAATAAGCTGTTGGTGAGAGTCTGCGAGCTTCGGGGAAGATATATTCGACTC 494
DB 421 AGTGCAATAAGCTGTTGGTGAGAGTCTGCGAGCTTCGGGGAAGATATATTCGACTC 480
QY 495 TGT'CAGAAATATTACTCTCCTCAGAACCCCGAGCAGTAGACTTCTTAGAATGTGCAAGAA 554
DB 481 TGT'CAGAAATATTACTCTCCTCAGAACCCCGAGCAGTAGACTTCTTAGAATGTGCAAGAA 540
QY 555 GCTAGAAAAAGATTAAATTCCTGGTCAAGACTCAAAACCAAGGCAAAATCCCAACTTG 614
DB 541 GCTAGAAAAAGATTAAATTCCTGGTCAAGACTCAAAACCAAGGCAAAATCCCAACTTG 600
QY 615 TTACCTGAAAGTTTCTGTAGATGGGGATACCAGAGATGGTCTCTGGTGAATGCTGTACTTC 674
DB 601 TTACCTGAAAGTTTCTGTAGATGGGGATACCAGAGATGGTCTCTGGTGAATGCTGTACTTC 660
QY 675 AAAGGAAGTGGAAAACTCCATTTGAGAAGAACTAAATGGGCTTTATCCCTTCGCTGTA 734
DB 661 AAAGGAAGTGGAAAACTCCATTTGAGAAGAACTAAATGGGCTTTATCCCTTCGCTGTA 720
QY 735 AACTCGGCTCAGCGCACACCTGTACAGATGATGTAATGCTGCTGAAAGCAATTTGGA 794
DB 721 AACTCGGCTCAGCGCACACCTGTACAGATGATGTAATGCTGCTGAAAGCAATTTGGA 780
QY 795 TACATAGAGACCTTAAGGCTCAGATTCTAGAACTCCCATATGCTGGAGATGTTAGCATG 854
DB 781 TACATAGAGACCTTAAGGCTCAGATTCTAGAACTCCCATATGCTGGAGATGTTAGCATG 840
QY 855 TTCTTGTGCTTCCAGATGAAATTCGCGATGTGTCTCACTGGCTGGAGCTGCTGAAAGT 914
DB 841 TTCTTGTGCTTCCAGATGAAATTCGCGATGTGTCTCACTGGCTGGAGCTGCTGAAAGT 900
QY 915 GAAATAACCTATGACAAACTCAACAGTGGACAGCAAGCAAAATGGCTGAAAGTGA 974
DB 901 GAAATAACCTATGACAAACTCAACAGTGGACAGCAAGCAAAATGGCTGAAAGTGA 960
QY 975 GTTGAGGTATACATACCCAGTTCAAAATAGAGAGCATTTAGAACTCAGATCCATTCTG 1034
DB 961 GTTGAGGTATACATACCCAGTTCAAAATAGAGAGCATTTAGAACTCAGATCCATTCTG 1020
QY 1035 AAAAGCATGGCATGGAGGACGCTTCAACAAGGAGCGGCAATTTCTCAGGAGTGTG 1094

b	1021	AGAAGCAATGGGCATGGAGGACGCCCTTCCAAAGAGGACGGGCCAAATTTCTCAGGGATGTCTG	1088
y	1095	GAGAGGAATGACCTGTTCTTTCTCTGAAGTGTTCACCAAGCCATGGTGGATGTGAATGAG	1154
b	1081	GAGAGGAATGACCTGTTCTTTCTCTGAAGTGTTCACCAAGCCATGGTGGATGTGAATGAG	1140
y	1155	GAGGCCACTGAAGCAGCGCTGGCACAGGAGTGTATGACAGGGAGAACTGCACATGGA	1214
b	1141	GAGGCCACTGAAGCAGCGCTGGCACAGGAGTGTATGACAGGGAGAACTGCACATGGA	1200
y	1215	GGCCACAGATTTGTGGCAGATCATCCGTTTCTTTTTCTTTATTATGATGAAGTAACCAAG	1274
b	1201	GGCCACAGTTTGTGGCAGATCATCCCTTTCTTTTCTTTATTATGATGAAGTAACCAAC	1260
y	1275	TGCATTTTATTTTTGGCGAGATTTTGGTCACACCTAAACCTAAGCGTCTCTGCAAA	1334
b	1361	TGCATTTTATTTTTGGCGAGATTTTCCACCCCTAAACTAAGCGTCTCTGCAAA	1320
y	1335	AGATTTTTGTAGATGAGCTGTGCGCTCAGAAATGCTATTTCAAATGGCAAAATTTAG	1394
b	1321	AGATTTTTGTAGATGAGCTGTGCGCTCAGAAATGCTATTTCAAATGGCAAAATTTAG	1380
y	1395	AGATGTTTTCTACATATTTCTGCTCTTCTGGAACAATCTGCTACCCACTAAATAAAAC	1454
b	1381	AGATGTTTTCTACATATTTCTGCTCTTCTGGAACAATCTGCTACCCACTAAATAAAAC	1440
y	1455	ACAGAAATAATPAGACAAATGCTATTATAACATGCAACCCCTATTAAATCATTTGTCCT	1514
b	1441	ACAGAAATAATPAGACAAATGCTATTATAACATGCAACCCCTATTAAATCATTTGTCCT	1500
y	1515	CTAAATGGGATCATGCCCAATTTAGATTTTTCCCTACTATCAGTTATTTTATAACATTA	1574
b	1501	CTAAATGGGATCATGCCCAATTTAGATTTTTCCCTACTATCAGTTATTTTATAACATTA	1560
y	1575	ACTTTTACTTTGTTATTTATTTATTTATAATGTTGAGTTTTTAAATTTATGCTCACTG	1634
b	1561	ACTTTTACTTTGTTATTTATTTATTTATAATGTTGAGTTTTTAAATTTATGCTCACTG	1620
y	1635	CCTATTTTAATGTAGCTAATAAGTTATAGACAGATGATCTGTTAATTTCCCTATCTAAT	1694
b	1621	CCTATTTTAATGTAGCTAATAAGTTATAGACAGATGATCTGTTAATTTCCCTATCTAAT	1680
y	1695	AAATGCCCTTTAATTTGTTCTCATATAAGAAATAGTAGSPATCCCTCCATGCCCTTCTGT	1754
b	1681	AAATGCCCTTTAATTTGTTCTCATATAAGAAATAGTAGSPATCCCTCCATGCCCTTCTGT	1740
y	1755	AATAAATATCTGGAAAAACATTAACATAGGCAATATATGCTTATGTCGATTTCTCTAGA	1814
b	1741	AATAAATATCTGGAAAAACATTAACATAGGCAATATATGCTTATGTCGATTTCTCTAGA	1800
y	1815	AATACATAACACATATATGCTGCTATCTTATTTCAATTCGAAGTATATAAT	1868
b	1801	ATAATACACACATATATGCTGCTATCTTATTTCAATTCGAAGTATATAAT	1854

RESULT 4
IS-07-693-636A-18

Sequence 18, Application US/076993636A
Patent No. 5426044
GENERAL INFORMATION:
APPLICANT: ANTALIS, Toni M.
APPLICANT: BARNES, Thomas M.
APPLICANT: CLARK, Michelle A.
APPLICANT: DEVINE, Peter L.
APPLICANT: GOSS, Neil H.
APPLICANT: LSHRBACH, Philip R.
TITLE OF INVENTION: RECOMBINANT PRODUCT
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
City: Alexandria
STATE: Virginia

COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/693,636A
FILING DATE: 19910430
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/025,815
FILING DATE: 13-MAR-1987
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/155 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2409 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA to genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 49..1296
US-07-693-636A-18

Query Match	96.8%;	Score 1847.6;	DB 1;	Length 2409;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1850; Conservative	0;	Mismatches	4;	Indels 0; Gaps 0
QY	15	GTCCAGACGCAACTCAGAGAATACCAGAGAAACAACCCAGATTGAAAACAATGGAGGATCTT	74	
Db	1	GTCCAGACGCAACTCAGAGAATAACCCAGAGAAACAACCCAGATTGAAAACAATGGAGGATCTT	60	
QY	75	TGTGTGGCAAACACACTCTTTTGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCCCC	134	
Db	61	TGTGTGGCAAACACACTCTTTTGCCCTCAATTTATTCAAGCATCTGGCAAAAGCAAGCCCC	120	
QY	135	ACCAGAACTCTTCCTCTCCCCATGGAGCATCTCGTCCACCATGGSCCATGGTCTACATG	194	
Db	121	ACCAGAACTCTTCCTCTCCCCATGGAGCATCTCGTCCACCATGGTCTACATG	180	
QY	195	GGTCTCAGGGGACGCCAGNAGCACCATGGCCAAAGTGCTTCAGTTTAAATGAAGTGGGA	254	
Db	181	GGTCTCAGGGGACGCCAAGAACCATAGATGGCCAAGTGCTTCAGTTTAAATGAAGTGGGA	240	
QY	255	GCCAAATGCAGTTACCCCCCATGACTCCAGAGAACTTTACCAAGCTGTGGTTTCATGCAGAG	314	
Db	241	GCCAAATGCAGTTACCCCCCATGACTCCAGAGAACTTTACCAAGCTGTGGTTTCATGCAGAG	300	
QY	315	ATCCAGNAGGTAAGTTATCTGTATGGATTTTCGAGGCACAAAGCTGCAGATAAAATCCAT	374	
Db	301	ATCCAGNAGGTAAGTTATCTGTATGGATTTTCGAGGCACAAAGCTGCAGATAAAATCCAT	360	
QY	375	TCATCTCTCCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGAATATTTTACTGGAA	434	
Db	361	TCATCTCTCCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGAATATTTTACTGGAA	420	
QY	435	AGTGTCAATAAGCTGTTTTGTGTGAGAGCTCTGCGAGCTTCGGGAAGAAATATATTTCGATC	494	
Db	421	AGTGTCAATAAGCTGTTTTGTGTGAGAGCTCTGCGAGCTTCGGGAAGAAATATATTTCGATC	480	
QY	495	TGTCAGAAAAATTACTCCTCAGAAACCCCAGCAGTAGACTTCTTAGAATGTGCGAAGAA	554	
Db	481	TGTCAGAAAAATTACTCCTCAGAAACCCCAGCAGTAGACTTCTTAGAATGTGCGAAGAA	540	

555 GCTAGAAAAAGATTAAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAACTTG 614
541 GCTAGAAAAAGATTAAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAACTTG 600
615 TTACCTGAAGGTTCTGTAGATGGGATACCAAGATGGTCTCTGGTGAATGCTGTACTTC 674
601 TTACCTGAAGGTTCTGTAGATGGGATACCAAGATGGTCTCTGGTGAATGCTGTACTTC 660
675 AAAGGAAAGTGGAAACCTCAATTTGAGAGAACTAAATGAGGCTTTATCCCTTCCTGGTGA 734
661 AAAGGAAAGTGGAAACCTCAATTTGAGAGAACTAAATGAGGCTTTATCCCTTCCTGGTGA 720
735 AACTCGGCTCAGCGACACCTGTACAGATGATGTAATGCGTGAAGCTTAACATTTGGA 794
721 AACTCGGCTCAGCGACACCTGTACAGATGATGTAATGCGTGAAGCTTAACATTTGGA 780
795 TACATAGAAAGCTTAAAGGCTCAGATCTTAGAACTCCCATATCTCGAGATGTTAGCATG 854
781 TACATAGAAAGCTTAAAGGCTCAGATCTTAGAACTCCCATATCTCGAGATGTTAGCATG 840
855 TTCTTGTTGCTTCCAGATGAAATTCGGATGTCCTCACTGGCTGGCTGGCTGGAAGT 914
841 TTCTTGTTGCTTCCAGATGAAATTCGGATGTCCTCACTGGCTGGCTGGAAGT 900
915 GAAATPAACTATGACAAACTCAACAACTGAGACCAAGCAAGCAAAATGGCTGAAGATGAA 974
901 GAAATPAACTATGACAAACTCAACAACTGAGACCAAGCAAGCAAAATGGCTGAAGATGAA 960
975 GTTGAGGTATACATACCCAGTTCAAATTAGAGAGCAATATGAATCAGATCCATCTG 1034
961 GTTGAGGTATACATACCCAGTTCAAATTAGAGAGCAATATGAATCAGATCCATCTG 1020
1035 AAAGCATGGGATGAGAGCGCTTCAACAAAGGCGGCGCAATTTCTCAGGATGTCG 1094
1021 AGAAGCATGGGATGAGAGCGCTTCAACAAAGGCGGCGCAATTTCTCAGGATGTCG 1080
1095 GAGAGGATGACTGTTCTTCTGAGGTGTTCCACCAAGCCATGGTGAATGAAATGAG 1154
1081 GAGAGGATGACTGTTCTTCTGAGGTGTTCCACCAAGCCATGGTGAATGAAATGAG 1140
1155 GAGGCGATGAGAGCGCTTCAACAAAGGCGGCGCAATTTCTCAGGATGTCG 1214
1141 GAGGCGATGAGAGCGCTTCAACAAAGGCGGCGCAATTTCTCAGGATGTCG 1200
1215 GCGCCACAGTTTGTGCGAGATCATCGTTCCTTTCTTATTATGCAATGATTAACCAAG 1274
1201 GCGCCACAGTTTGTGCGAGATCATCGTTCCTTTCTTATTATGCAATGATTAACCAAG 1260
1275 TGCAATTTATTTTTCGCGAGATTTGCTCAGCACTTCAACAAAGGCGGCTGCTGCAAA 1334
1261 TGCAATTTATTTTTCGCGAGATTTCTCAGCACTTCAACAAAGGCGGCTGCTGCAAA 1320
1335 AGATTTTGTAGATGAGCTGTGCTCAGATGCTTCAATTTCAATTTGCCAAATTTAG 1394
1321 AGATTTTGTAGATGAGCTGTGCTCAGATGCTTCAATTTCAATTTGCCAAATTTAG 1380
1395 AGATGTTTCTACATATTTGCTGCTCAGCACTTCAACAAAGGCGGCTGCTGCAAA 1454
1381 AGATGTTTCTACATATTTGCTGCTCAGCACTTCAACAAAGGCGGCTGCTGCAAA 1440
1455 ACAGAAATTAATAGCAATTTGCTTATTAACATGACAAACCTTATTAATCAATTTGGTCTT 1514
1441 ACAGAAATTAATAGCAATTTGCTTATTAACATGACAAACCTTATTAATCAATTTGGTCTT 1500
1515 CTAATAATGGATCATGCCATTTAGATTTTCTTACTATCAATTTATTTTATTAACATTA 1574
1501 CTAATAATGGATCATGCCATTTAGATTTTCTTACTATCAATTTATTTTATTAACATTA 1560
1575 ACTTTTACTTTGTTATTTATTTATTTATTAATGAGTGTGTTTAAATTTGCTCACTG 1634
1561 ACTTTTACTTTGTTATTTATTTATTTATTAATGAGTGTGTTTAAATTTGCTCACTG 1620

1635 CCTATTTAATGTAGCTAATAAGTTATAGAACAGATGATCTGTTAATTTCTTATCTAAT 1694
1621 CCTATTTAATGTAGCTAATAAGTTATAGAACAGATGATCTGTTAATTTCTTATCTAAT 1680
1695 AAATGCCCTTTAATTTCTTCATTAATGAAGAAATAGTAGTATCCCTCCATGCCCTTCTGT 1754
1681 AAATGCCCTTTAATTTCTTCATTAATGAAGAAATAGTAGTATCCCTCCATGCCCTTCTGT 1740
1755 AATAAATATCTGGAATAAATCAATTAACATAGGCAAAATATATGTTATGTCATTTCTAGA 1814
1741 AATAAATATCTGGAATAAATCAATTAACATAGGCAAAATATATGTTATGTCATTTCTAGA 1800
1815 AATACATACACATATATGTCGTATCTTATTTCAATTTGCAAGTATATAAT 1868
1801 AATACATACACATATATGTCGTATCTGTTATTTCAATTTGCAAGTATATAAT 1854

RESULT 5
US-07-768-286B-1
; Sequence 1, Application US/07768286B
; Patent No. 5444153
; GENERAL INFORMATION:
; APPLICANT: GOSS, Neil H.
; APPLICANT: RICHARDSON, Michael A.
; TITLE OF INVENTION: VARIANTS OF PAI-2
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/768,286B
; FILING DATE: 19911011
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00603
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/157 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1610 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL TYPE: Monocyte
; CELL LINE: U937
; IMMEDIATE SOURCE:
; CLONE: BTA 1445
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 18
; MAP POSITION: 18q21-q23
; FEATURE:
; NAME/KEY: CDS

LOCATION: 49..1296									
OTHER INFORMATION: /product= "human plasminogen									
OTHER INFORMATION: activator inhibitor type 2 protein"									
IS-07-768-286B-1									
Query Match									
Best Local Similarity 84.0%; Score 1602; DB 1; Length 1610;									
Matches 1605; Conservative 0; Mismatches 5; Indels 0; Gaps 0;									
15	GT	CAGACAGCAACTCAGAGATAAACCAGAGAACCAACAGATTGAAACAATGGAGGATCTT	74						
16	GT	CAGACAGCAACTCAGAGATAAACCAGAGAACCAACAGATTGAAACAATGGAGGATCTT	60						
75	TG	TGTCGCAACACACTCTTTCCTCTCAATTTATTCAGCATCTGCGCAAAAGCAAGCC	134						
61	TG	TGTCGCAACACACTCTTTCCTCTCAATTTATTCAGCATCTGCGCAAAAGCAAGCC	120						
135	ACC	CAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGTCTACATG	194						
121	ACC	CAGAACCTCTTCTCTCCCATGGAGCATCTCGTCCACCATGGCCATGGTCTACATG	180						
195	GG	CTCCAGGGGCGAGCAGCAGAGACAGATGGCAAGGTCTTCAGTTTAAATGAAGTGGGA	254						
181	GG	CTCCAGGGGCGAGCAGCAGAGACAGATGGCCAGGTCTTCAGTTTAAATGAAGTGGGA	240						
255	GC	CAATGCAATTACCCCATGATCTCAGAGAACTTTACAGCTGTGGTTCATGCGAGCAG	314						
241	GC	CAATGCAATTACCCCATGATCTCAGAGAACTTTACAGCTGTGGTTCATGCGAGCAG	300						
315	AT	CAGAGGGGTAGTTATCTCTGATCGAATTTGCGAGCACAAGCTGCAGATAAAATCCAT	374						
301	AT	CAGAGGGGTAGTTATCTCTGATCGAATTTGCGAGCACAAGCTGCAGATAAAATCCAT	360						
375	TC	ATCTCCCTCTCTCAGCTCTCAATCAATGATCCACAGGGAATTTATTTACTGGAA	434						
361	TC	ATCTCCCTCTCTCAGCTCTCAATCAATGATCCACAGGGAATTTATTTACTGGAA	420						
435	AG	TCATAAAGCTTTTGGTGAAGTCTGGAGCTTCGGGAAGAAATATATTCGACTC	494						
421	AG	TCATAAAGCTTTTGGTGAAGTCTGGAGCTTCGGGAAGAAATATATTCGACTC	480						
495	TG	CAGAAATATTACTCTCAGAACCCAGGAGTAGACTTCCTAGATGTGAGAGAGAA	554						
481	TG	CAGAAATATTACTCTCAGAACCCAGGAGTAGACTTCCTAGATGTGAGAGAGAA	540						
555	GCT	AGAAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAACTTG	614						
541	GCT	AGAAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAACTTG	600						
615	TT	ACTCTGAAGGTTCTGTAGATGGGATACCAGGATGGTCTGTGATGTCTACTTC	674						
601	TT	ACTCTGAAGGTTCTGTAGATGGGATACCAGGATGGTCTGTGATGTCTACTTC	660						
675	AA	GGAAAGTGAAGAACTCCATTTGAGAGAACTAAATGGCCCTTATCCTTTCCGTTA	734						
661	AA	GGAAAGTGAAGAACTCCATTTGAGAGAACTAAATGGCCCTTATCCTTTCCGTTA	720						
735	AA	CTCGGCTCAGCGCACACTGTACAGATGATGATCTTTGCGTGAAGAGCTAAACATGGA	794						
721	AA	CTCGGCTCAGCGCACACTGTACAGATGATGATCTTTGCGTGAAGAGCTAAACATGGA	780						
795	TAC	ATAGAGACCTAAAGGCTCAGATTCCTAGAGATCCCATATGCTGGAGATGTAGCATG	854						
781	TAC	ATAGAGACCTAAAGGCTCAGATTCCTAGAGATCCCATATGCTGGAGATGTAGCATG	840						
855	TT	CTTGTCTCCAGATGAATTCGGATGTGTCCATGGCTGGAGCTGGGAAAGT	914						
841	TT	CTTGTCTCCAGATGAATTCGGATGTGTCCATGGCTGGAGCTGGGAAAGT	900						
915	GAA	ATAACCTATGACAAACTCAACAAGTGGACCAAGACAAATAATGGCTGAAGATGAA	974						
901	GAA	ATAACCTATGACAAACTCAACAAGTGGACCAAGACAAATAATGGCTGAAGATGAA	960						

QY	975	GTT	GAGGTATACATACCCAGTTCAAATTAGAAGAGCATTATGAATCAGATCCATTCG	1034
DB	961	GTT	GAGGTATACATACCCAGTTCAAATTAGAAGAGCATTATGAATCAGATCCATTCG	1020
QY	1035	AA	AGCATGGCATGGAGGAGCCCTTCAACAGGAGCGGCGCAATTTCTCAGGATGTCG	1094
DB	1021	AGA	AGCATGGCATGGAGGAGCCCTTCAACAGGAGCGGCGCAATTTCTCAGGATGTCG	1080
QY	1095	GAG	AGGAATACCTGTTCTTCTGAAAGTGTCCACCAAGCCATGGTGGATGAATGAG	1154
DB	1081	GAG	AGGAATACCTGTTCTTCTGAAAGTGTCCACCAAGCCATGGTGGATGAATGAG	1140
QY	1155	GAG	GCGCATGAAGCAGCGCGCTGGCACAGAGGTGTATGACAGGGAGAACTGACATGGA	1214
DB	1141	GAG	GCGCATGAAGCAGCGCGCTGGCACAGAGGTGTATGACAGGGAGAACTGACATGGA	1200
QY	1215	GG	CCACAGTTTGGGAGATCATCGTTTCTTTCTTTATGATGATAGATAACCAAG	1274
DB	1201	GG	CCACAGTTTGGGAGATCATCGTTTCTTTCTTTATGATGATAGATAACCAAC	1260
QY	1275	TGC	ATTTTATTTTGGGAGATTTTGTCTCACCCCTAAACCTAAGCGTCTCTTCTGCAAA	1334
DB	1261	TGC	ATTTTATTTTGGGAGATTTTCTCTCACCCCTAAACCTAAGCGTCTCTTCTGCAAA	1320
QY	1335	AGA	TTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTCAAATTTGCCAAAATTTAG	1394
DB	1321	AGA	TTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTCAAATTTGCCAAAATTTAG	1380
QY	1395	AGA	TTTTGTACATATTTTGTCTCTTCTGAAACAACCTCTGCTACCCATAAATAAAC	1454
DB	1381	AGA	TTTTGTACATATTTTGTCTCTTCTGAAACAACCTCTGCTACCCATAAATAAAC	1440
QY	1455	ACA	GAATAAATAGACAATTTGCTATTATAACATGACAAACCTTATTAATCATTTGGTCTT	1514
DB	1441	ACA	GAATAAATAGACAATTTGCTATTATAACATGACAAACCTTATTAATCATTTGGTCTT	1500
QY	1515	CT	AAATGGGATCATGCCCATTTAGATTTTCTTCTACTATCAGTTATTTTATAACATTA	1574
DB	1501	CT	AAATGGGATCATGCCCATTTAGATTTTCTTCTACTATCAGTTATTTTATAACATTA	1560
QY	1575	ACT	TTTACTTTGTTATTTATTTATTAATAATGTTGAGTTTTTAATAA	1624
DB	1561	ACT	TTTACTTTGTTATTTATTTATTAATAATGTTGAGTTTTTAATAA	1610

RESULT 6
US-07-768-286B-3
; Sequence 3, Application US/07768286B
; Patent No. 544153
; GENERAL INFORMATION:
; APPLICANT: GOSS, Neil H.
; APPLICANT: RICHARDSON, Michael A.
; TITLE OF INVENTION: VARIANTS OF PAI-2
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/768,286B
; FILING DATE: 19911011
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00603
; FILING DATE: 20-DEC-1990

ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/157 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1512 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: Monocyte
CELL LINE: U937
IMMEDIATE SOURCE:
CLONE: BTA 1916
FEATURE:
NAME/KEY: CDS
LOCATION: 22..1200
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "Product binds to
OTHER INFORMATION: urokinase, tissue plasminogen activator"
OTHER INFORMATION: /product= "PAI-2 variant, protease sensitive site
OTHER INFORMATION: removed"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /notes= "Codes for human plasminogen activator
OTHER INFORMATION: inhibitor type 2 protein in which amino acids 74
OTHER INFORMATION: to 96 inclusive have been deleted."
US-07-768-286B-3

Query Match 73.5%; Score 1402.6; DB 1; Length 1512;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 1484; Conservative 0; Mismatches 4; Indels 69; Gaps 1;

Qy	62	AATGAGGATCTTTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGGC	121
Db	21	AATGAGGATCTTTGTGGCAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTGGC	80
Qy	122	AAAGCAAGCCCAACCCAGAACCTCTCTCTCCATGGAGCATCTCGTCCACCATGGC	181
Db	81	AAAGCAAGCCCAACCCAGAACCTCTCTCTCCATGGAGCATCTCGTCCACCATGGC	140
Qy	182	CATGGTCTACATGGGTCGAGGGGAGCAGCAGAGACAGATGGCCAGGTGCTTCAGTT	241
Db	141	CATGGTCTACATGGGTCGAGGGGAGCAGCAGAGACAGATGGCCAGGTGCTTCAGTT	200
Qy	242	TAATGAAGTGGAGCCAAATGCAGTTACCCCATGACTCCAGAGAACTTTACCAGCTGTGG	301
Db	201	TAATGAAGTGGAGCCAAATGCAGTTACCCCATGACTCCA-----	240
Qy	302	GTTTCAGCAGCATCCAGAGGGTAGTTATCTCTGATGCGATTTTGGAGGCAGAGCTGC	361
Db	241	-----GCACAGCTGC	251
Qy	362	AGATAAAATCCATTATCTCTCTCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGAA	421
Db	252	AGATAAAATCCATTATCTCTCTCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGAA	311
Qy	422	TTATTTACTGGAAGTGTCAATAGCTGTTGGTGAGAGCTCGGAGCTTCCGGGAAGA	481
Db	312	TTATTTACTGGAAGTGTCAATAGCTGTTGGTGAGAGCTCGGAGCTTCCGGGAAGA	371
Qy	482	ATATATTGCACTCTGTGAGAAATATTACTCTCAGAACCCAGGAGTAGACTTCCCTAGA	541
Db	372	ATATATTGCACTCTGTGAGAAATATTACTCTCAGAACCCAGGAGTAGACTTCCCTAGA	431

Qy	542	ATGTCAGAGAGCTAGAAAAAGATTAAATCTCTGGTCAAGACTCAAAACCAAGGCAA	601
Db	432	ATGTCAGAGAGCTAGAAAAAGATTAAATCTCTGGTCAAGACTCAAAACCAAGGCAA	491
Qy	602	AATCCAAACTTTGTTACCTGAAGTTCTGTAGATGGGATACCAAGGATGTCCTGGTCAA	661
Db	492	AATCCAAACTTTGTTACCTGAAGTTCTGTAGATGGGATACCAAGGATGTCCTGGTCAA	551
Qy	662	TGCTGTCTACTTCAAGGAAAGTGGAAAACTCCATTTGAGAAGAACTAAATGGGCTTTA	721
Db	552	TGCTGTCTACTTCAAGGAAAGTGGAAAACTCCATTTGAGAAGAACTAAATGGGCTTTA	611
Qy	722	TCCTTTCCGTGTAAACTCGGCTCAGCGCACCTGTACAGATGATGATCTTCGTGAAA	781
Db	612	TCCTTTCCGTGTAAACTCGGCTCAGCGCACCTGTACAGATGATGATCTTCGTGAAA	671
Qy	782	GCTAAACATTTGGATACATAGAAGACCTAAAGGCTCAGATTCTAGAACTCCCATGCTGG	841
Db	672	GCTAAACATTTGGATACATAGAAGACCTAAAGGCTCAGATTCTAGAACTCCCATGCTGG	731
Qy	842	AGATGTTAGCATGTTCTTGTGCTTCCAGATGAAATGCGCATGTGTCCACTGGCTTGA	901
Db	732	AGATGTTAGCATGTTCTTGTGCTTCCAGATGAAATGCGCATGTGTCCACTGGCTTGA	791
Qy	902	GCTGCTGAAAGTGAATAAACCCTATGACAACTCAACAAGTGGACCAAGACAAAT	961
Db	792	GCTGCTGAAAGTGAATAAACCCTATGACAACTCAACAAGTGGACCAAGACAAAT	851
Qy	962	GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGACATTTAGAACT	1021
Db	852	GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGACATTTAGAACT	911
Qy	1022	CAGATCCATCTGAAAGATGGCGATGGAGGCGCTTCAACAGGAGGAGGCGGCAATTT	1081
Db	912	CAGATCCATCTGAGAGGATGGCGATGGAGGCGCTTCAACAGGAGGAGGCGGCAATTT	971
Qy	1082	CTCAGGGATGTCGGAGAGGAATGACCTGTTCTTCTGAAAGTGTTCACCAAGCCATGGT	1141
Db	972	CTCAGGGATGTCGGAGAGGAATGACCTGTTCTTCTGAAAGTGTTCACCAAGCCATGGT	1031
Qy	1142	GGATGTGAATGAGGAGGCGACTGAAGCAGCGCTGGCAGCAGGAGGTGTATGACAGGAG	1201
Db	1032	GGATGTGAATGAGGAGGCGACTGAAGCAGCGCTGGCAGCAGGAGGTGTATGACAGGAG	1091
Qy	1202	AACTGGACATGGAGGCGCCAGTTTGTGGAGATCATCCGTTCTTTTCTTTTATATGCA	1261
Db	1092	AACTGGACATGGAGGCGCCAGTTTGTGGAGATCATCCCTTTCTTTTCTTTTATATGCA	1151
Qy	1262	TAAGTAAACCAAGTGCATTTTATTTTTCGCGAGATTTTGTCTCACCTAAAGCGGTG	1321
Db	1152	TAAGTAAACCAAGTGCATTTTATTTTTCGCGAGATTTTGTCTCACCTAAAGCGGTG	1211
Qy	1322	CTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTCAAAT	1381
Db	1212	CTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTCAAAT	1271
Qy	1382	GCCAAAAATTAGAGATGTTTCTTACATATTTCTGCTCTTCTGAAACAATCTCTGCTACC	1441
Db	1272	GCCAAAAATTAGAGATGTTTCTTACATATTTCTGCTCTTCTGAAACAATCTCTGCTACC	1331
Qy	1442	ACTAAATAAAACACAGAAATAATAGACAAATGCTCTATTAATACATGACACCTCTTA	1501
Db	1332	ACTAAATAAAACACAGAAATAATAGACAAATGCTCTATTAATACATGACACCTCTTA	1391
Qy	1502	ATCATTTGGTCTTCTTAAATGGGATCATGCCAATTTAGATTTTCTTACTATCAGTTTAT	1561
Db	1392	ATCATTTGGTCTTCTTAAATGGGATCATGCCAATTTAGATTTTCTTACTATCAGTTTAT	1451
Qy	1562	TTTATAACATTAACCTTTTACTTTTATTTATTTATTTATTTATATATGTTGAGTTTTT	1618
Db	1452	TTTATAACATTAACCTTTTACTTTTATTTATTTATTTATTTATATGTTGAGTTTTT	1508

```

RESULT 7
IS-07-768-286B-5
Sequence 5, Application US/07768286B
Patent No. 544153
GENERAL INFORMATION:
APPLICANT: GOSS, Neil H.
APPLICANT: RICHARDSON, Michael A.
TITLE OF INVENTION: VARIANTS OF PAI-2
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/768,286B
FILING DATE: 19911011
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00603
FILING DATE: 20-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/157 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: Monocyte
CELL LINE: U937
IMMEDIATE SOURCE:
CLONE: BTA 1922
FEATURE:
NAME/KEY: CDS
LOCATION: 22...1170
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "Product binds to
OTHER INFORMATION: /product= "PAI-2 variant, protease sensitive site
OTHER INFORMATION: /removed=
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /note= "Codes for human plasminogen activator
OTHER INFORMATION: inhibitor type 2 protein in which amino acids 66
OTHER INFORMATION: to 98 inclusive have been deleted."
US-07-768-286B-5
Query Match 70.4%; Score 1342.6; DB 1; Length 1482;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 1454; Conservative 0; Mismatches 4; Indels 99; Gaps 1;
QY 62 AATGAGGATCTTTGTGGGAAAACACACTCTTTTGGCCCTCAATTATTCAGGATCTGGC 121
DB 21 AATGAGGATCTTTGTGGGAAAACACACTCTTTTGGCCCTCAATTATTCAGGATCTGGC 80

```

[illegible][illegible]

970 TGGCTGAGATGAAGTTGAGGTATACATACCCAGTTCAAATTGAGAGCATTATGAAC 1029
1021 TCAGATCCATTCTGAAAGCATGGGCATGAGGACGCTTCAACAGGAGACGGGCCAATT 1080
1030 TCAGATCCATTCTGAGAGCATGGGCATGAGGACGCTTCAACAGGAGACGGGCCAATT 1089
1081 TCTCAGGGATGTCGAGAGGAGTACCTGTTTCTTTCTGAGGTGTTCCACCAAGCCATGG 1140
1090 TCTCAGGGATGTCGAGAGGAGTACCTGTTTCTTTCTGAGGTGTTCCACCAAGCCATGG 1149
1141 TGGATGTGAATGAGAGGACCTGAGCAGCCGCTGGGCACAGAGGTGTTATGACAGGGA 1200
1150 TGGATGTGAATGAGAGGACCTGAGCAGCCGCTGGGCACAGAGGTGTTATGACAGGGA 1209
1201 GAACCTGACATGGAGGCCCCACAGTTGTGGCAGATCATCCGTTTCTTTTCTTATTATGC 1260
1210 GAACCTGACATGGAGGCCCCACAGTTGTGGCAGATCATCCGTTTCTTTTCTTATTATGC 1269
1261 ATAAGATAACCAAGTGCATTTATTTTTCGGCAGATTTTGCCTACCCCTAAACTAAGCG 1319
1270 ATAAGATAACCAAGTGCATTTATTTTTCGGCAGATTTTTCCTACCCCTAAACTAAGCG 1328

RESULT 9

JS-07-679-052A-16
Sequence 16, Application US/07679052A
Patent No. 5298400
GENERAL INFORMATION:
APPLICANT: WHITFIELD, Peter L.
APPLICANT: RICHARDSON, Michael A.
APPLICANT: BUNN, Clive L.
TITLE OF INVENTION: RECOMBINANT PRODUCT
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/679,052A
FILING DATE: 19910506
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/147 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1340 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid;
DESCRIPTION: Synthetic DNA oligonucleotide
FEATURE:
NAME/KEY: CDS
LOCATION: 15...1331
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 15...86
OTHER INFORMATION: /product= "a-1-antitrypsin signal"
US-07-679-052A-16

Query Match 65.3%; Score 1246; DB 1; Length 1340;
Best Local Similarity 99.6%; Pred. No. 2e-292; 5; Indels 0; Gaps 0;
Matches 1249; Conservative 0; Mismatches 0;
Qy 66 GAGGATCTTTGTGGCAAAACACACTCTTTTGGCCCTCAATTTATTCAAGCATCTGGCAAAA 125
Db 87 GAAGATCTTTGTGGCAAAACACACTCTTTTGGCCCTCAATTTATTCAAGCATCTGGCAAAA 146
Qy 126 GCAAGCCCAACCCAGAACCTCTTCTCTCCCATGAGCATCTCGTCCACCATGCGCATG 185
Db 147 GCAAGCCCAACCCAGAACCTCTTCTCTCCCATGAGCATCTCGTCCACCATGCGCATG 206
Qy 186 GTCTACATGGGCTCCAGGGGAGACCGAAGACCAAGATGSCCAAGGTGCTTCAGTTTAA 245
Db 207 GTCTACATGGGCTCCAGGGGAGACCGAAGACCAAGATGSCCAAGGTGCTTCAGTTTAA 266
Qy 246 GAAGTGGAGCAATGTCAGTTTACCCCATGACTCCAGAGAACTTTACAGCTGTGGGTTC 305
Db 267 GAAGTGGAGCAATGTCAGTTTACCCCATGACTCCAGAGAACTTTACAGCTGTGGGTTC 326
Qy 306 ATGACAGATCCAGAAAGGTAGTTATCTGTATGCGATTTTGACGACCAAGCTGCAGAT 365
Db 327 ATGACAGATCCAGAAAGGTAGTTATCTGTATGCGATTTTGACGACCAAGCTGCAGAT 386
Qy 366 AAAATCCATTCACTCTTCCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGAATTAT 425
Db 387 AAAATCCATTCACTCTTCCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGAATTAT 446
Qy 426 TTACTGAAAGTGTCAATAAGCTGTTTGTGAGAGTCTGCGAGCTTCCGGGAGAAATAT 485
Db 447 TTACTGAAAGTGTCAATAAGCTGTTTGTGAGAGTCTGCGAGCTTCCGGGAGAAATAT 506
Qy 486 ATTGACTCTGTGAGAAATATTACTCTCTCAGAACCCAGGAGTAGACTTCTCTAGAATGT 545
Db 507 ATTGACTCTGTGAGAAATATTACTCTCTCAGAACCCAGGAGTAGACTTCTCTAGAATGT 566
Qy 546 GCAGAGAGCTAGAAAAAGATTAATCTCTGGTCAAGACTCAACCAAGGCAAAATC 605
Db 567 GCAGAGAGCTAGAAAAAGATTAATCTCTGGTCAAGACTCAACCAAGGCAAAATC 626
Qy 606 CCAACTTGTGTACCTCAAGGTTCTGTAGATGGGATACCAAGGATGGTCTTGGTGAATGCT 665
Db 627 CCAACTTGTGTACCTCAAGGTTCTGTAGATGGGATACCAAGGATGGTCTTGGTGAATGCT 686
Qy 666 GTCTACTTCAAGGAAAGTGGAAACTCCATTTGAGAGAACTTAATGGGCTTTATCCT 725
Db 687 GTCTACTTCAAGGAAAGTGGAAACTCCATTTGAGAGAACTTAATGGGCTTTATCCT 746
Qy 726 TTCCGTGTAACCTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGCCTGAAAAGCTA 785
Db 747 TTCCGTGTAACCTCGGCTCAGCGCACACCTGTACAGATGATGTACTTGCCTGAAAAGCTA 806
Qy 786 AACATTGGATACATAGAACCTTAAGGCTCAGATTCTAGAACTCCCATATGCTGGAGAT 845
Db 807 AACATTGGATACATAGAACCTTAAGGCTCAGATTCTAGAACTCCCATATGCTGGAGAT 866
Qy 846 GTTAGCATGTTCTTGTGCTTCCAGATGAAATTCGCGATGTGTCCTAGCTGGCTGGAGCTG 905
Db 867 GTTAGCATGTTCTTGTGCTTCCAGATGAAATTCGCGATGTGTCCTAGCTGGCTGGAGCTG 926
Qy 906 CTGGAAGTGAATAACCTATGACAACTCAACAGTGGACCAAGCAAAATGGCT 965
Db 927 CTGGAAGTGAATAACCTATGACAACTCAACAGTGGACCAAGCAAAATGGCT 986
Qy 966 GAAGATGAAGTTGAGGTATACATACCCAGTTCAAAATTAGAAGAGCATTTATGAATCAGA 1025
Db 987 GAAGATGAAGTTGAGGTATACATACCCAGTTCAAAATTAGAAGAGCATTTATGAATCAGA 1046
Qy 1026 TCCATTCTGAAAAGCATGGGAGGAGCGCTTCAACAGGAGCGGCAATTTCTCA 1085
Db 1047 TCCATTCTGAAAAGCATGGGAGGAGCGCTTCAACAGGAGCGGCAATTTCTCA 1106

QY 1086 GGGATGTCGGAGAGGATGACCTGTTCTTTCTGAAGTGTTCACCAAGCCATGGTGGAT 1145
 Db 1107 GGGATGTCGGAGAGGATGACCTGTTCTTTCTGAAGTGTTCACCAAGCCATGGTGGAT 1166
 QY 1146 GTGAATGAGGAGGCACTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGAGAACT 1205
 Db 1167 GTGAATGAGGAGGCACTGAAGCAGCGCTGGCAGAGGAGTGTATGACAGGAGAACT 1226
 QY 1206 GGACATGGAGGCCCCACAGTTGTGGCAGATCATCGTTTCTTTTCTTATATGATTAAG 1265
 Db 1227 GGACATGGAGGCCCCACAGTTGTGGCAGATCATCGTTTCTTTTCTTATATGATTAAG 1286
 QY 1266 ATACCAAGTGCATTTATTTTGGCAGATTTTGTCCACCCCTAAACTAAGCG 1319
 Db 1287 ATACCAAGTGCATTTATTTTGGCAGATTTTGTCCACCCCTAAACTAAGCG 1340

RESULT 10
 US-10-024-427-3
 ; Sequence 3, Application US/10024427
 ; Patent No. 6583269
 ; GENERAL INFORMATION:
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Nepomnichy, Boris
 ; APPLICANT: Walke, D. Wade
 ; TITLE OF INVENTION: No. 6583269el Human Protease Inhibitor and Polynucleotides Encodi
 ; FILE REFERENCE: LEX-0283-USA
 ; CURRENT APPLICATION NUMBER: US/10/024,427
 ; PRIOR APPLICATION NUMBER: 2001-12-18
 ; PRIOR FILING DATE: 2000-12-18
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1519
 ; TYPE: DNA
 ; ORGANISM: homo sapiens
 US-10-024-427-3

Query Match 14.6%; Score 279; DB 4; Length 1519;
 Best Local Similarity 53.6%; Pred. No. 4.4e-58;
 Matches 688; Conservative 0; Mismatches 565; Indels 30; Gaps 4;

QY 60 ACATGGAGATCTTGTGTGGGAAACACACTCTTTGGCCCTCAATTTATTCAAGCATCTG 119
 Db 142 ACAATGGACTCTCTGTACAGCAACACCAATTTTGTCTTGATCTTTTCAAGAGATA 201
 QY 120 GCAAAAGCAAGCCCCACCCAGAACCTCTCTCTCCCATGGAGCATCTCGTCCACCATG 179
 Db 202 GCAAAAGATGATGTCATATAAAACATAATTTTCTCTCCCTGAGCCTCTCAGTCGCCCT 261
 QY 180 GCATGGTCTACATGGCTCCAGGGCAGCACCGAAGACAGATGGCCAGGTGCTTCAG 239
 Db 262 GGTATGGTACGCTTGGGTGTAGAAAGTACAGTGCATCATCAGATTTGATGAGGTACTAC 321
 QY 240 TTTATGAAGTGGAGGCAATGCAATTA-----CCCCATGACTCCAGAGAACTTTA 291
 Db 322 TTCAACGAATTTTCCAGATGAAAGCAAGAACTGACCTGCTGTGAAAGCAACAAA 381
 QY 292 CNGCTGTGGTTCATCGAGC-AGATCAGAGGGTAGTTATCTGTGATCGGATTTGCGAG 350
 Db 392 CAAAAGTGTGGTGTGACGCTCTCTGGAGGGGCGAGAAAACGACAGAGCCTCTGGAT 441
 QY 351 GCACAGCTGCAGATAAAATCCAT-----TCATCCTCCGCTCTCTC 392
 Db 442 CAGCAGGCTGGTCTTAAACAAATGAGAGCGAGCTGGTTCAGTGTACTTTTGGGAGCTT 501
 QY 393 AGCTCTGCATCATGATCCACAGGGAATTTTACTGGAAAGTGTCAATAAGCTGTTT 452
 Db 502 CTCTCCAAATTAGACAGATCAAGACTGATTACACACTGAGTATTCACACAGGCTTAT 561
 QY 453 GGTGAGAGTCTCGAGCTTCCGGGAGAAATATATTCGACTGTGCGAATATTAATCTCC 512

Db 562 GGAGAGCAGGAATCCCAATCTGTCTCAGGATACTTAGATGGTGTGATTCATTTTACCAC 621
 QY 513 TCAGAACCCAGCAGTACATCTTCTAGATGTGTCAGAGAGCTAGAAAAAGATTAAT 572
 Db 622 ACGAGATTAAGAGTGTGATTTTCAAAAAAACCCCTGAAAAATCCAGACAGAGATTAA 681
 QY 573 TCCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAACTTGTTCACCTGAAAGTTCTGTA 632
 Db 682 TTCTGGTTGAATGTCAATCCCAAGGTAAATCAAGGAACCTTTTCAGCAAGGACGCTATT 741
 QY 633 GATGGGATACAGGATGTCTCTGGTGAATGCTGTCTACTTCAAGGAAGTGGAAACT 692
 Db 742 AATGCTGAGACTGTGTGGTACTGGTGAATGCTGTCTTCTTCAAGGCCAAATGGGAACA 801
 QY 693 CCATTTGAGAAGAACTAAATGGGCTTTATCTTTTCCGTGTAAACTCGGCTCAGCGACA 752
 Db 802 TACTTTGACCATGAAAAACACGGTGGATGCACCTTTCTGTCTAAATGCGAATGAAAAACA 861
 QY 753 CCTGTACAGATGATGTCTGCGTGAAGCTAAACATTTGGATACATAGAGACCTTAAG 812
 Db 862 AGTGAAGATGATGAGCAAAAGGACTCTACAGAAATTTGGCTTCATAGAGAGGTGAAG 921
 QY 813 GCTCAGATCTAGAACTCCCATATGCT---GGAGATGTAGCATGTTCTTTGTGCTTCCA 869
 Db 922 GCACAGATCTGGAATGAGGTACACCAAGGGGAAGCTCAGCATGTTCTGCTGTGCCA 981
 QY 870 GATGAATTCGCCATGTCTCACTGGCTTGGAGCTGTGGAAGTGAATAACCTATGAC 929
 Db 982 TCTCACTTAAGATTAACCTGAAGGCTCTGGAAGCTTGAAGGAAATCACTATGAA 1041
 QY 930 AAATCAACAGTGGACCAAGCAAAATGCTGAAGTGAAGTGAAGTGAAGTATACATA 989
 Db 1042 AAAATGGTGGCTGGAGCAGCTCAGAAAAACATGTGAGAAAGATCGTGTCTGCTCTTC 1101
 QY 990 CCCAGTTCAAAATTAGAAGACATTAATGAACTCAGATCCATTCTCAAAAGCATGGCATG 1049
 Db 1102 CCCGCTTACCTGGAGACAGTATGATCTCAATCCATTTTACAGACATGGCAAT 1161
 QY 1050 GAGGACCCCTTCAACAGGGACGGGCCAATTTCTCAGGGATGTGCGAGGAGTAACCTG 1109
 Db 1162 ACGGATATCTTTGATGAAACGAGGGCTGATCTTACTGGAATCTCTCCAAGTCCCAATTTG 1221
 QY 1110 TTTCTTTCTGAAGTGTTCACCAAGCCATGTGTGATGTAATGAGGAGGCACTGAAGCA 1169
 Db 1222 TACTTGTCAAAATTAATCCACAAACCTTTGTGGAGTGGATGAAAACGGTACCCAGGCA 1281
 QY 1170 GCGCTGGCAGAGAGTGTATGACAGGGAGAACTGGAAGTGAAGTGAAGTGAAGTGAAG 1229
 Db 1282 GCTGCAAGCACTGGGCTGTTGTCTCGAAAGGTCACTACGATCTTGGTGGAGTTAAT 1341
 QY 1230 GCAGATCATCCGTTTCTTTTCTTATTATGATAAGATAACCAAGTGCATTTATTTTTC 1289
 Db 1342 GCAACACCCCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTAT 1401
 QY 1290 GGAGATTTTGTCCACCCCTAAA 1312
 Db 1402 GCGAGGTCTGCTCTCTCTTAAA 1424

RESULT 11
 US-10-024-427-1
 ; Sequence 1, Application US/10024427
 ; Patent No. 6583269
 ; GENERAL INFORMATION:
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Nepomnichy, Boris
 ; APPLICANT: Walke, D. Wade
 ; TITLE OF INVENTION: No. 6583269el Human Protease Inhibitor and Polynucleotides Encodi
 ; FILE REFERENCE: LEX-0283-USA
 ; CURRENT APPLICATION NUMBER: US/10/024,427
 ; CURRENT FILING DATE: 2001-12-18
 ; PRIOR APPLICATION NUMBER: US 60/256,287
 ; PRIOR FILING DATE: 2000-12-18

NUMBER OF SEQ ID NOS: 3
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1
 LENGTH: 1278
 TYPE: DNA
 ORGANISM: homo sapiens
 IS-10-024-427-1

Query Match 14.4%; Score 274; DB 4; Length 1278;
 Best Local Similarity 53.4%; Pred. No. 6.6e-57;
 Matches 683; Conservative 0; Mismatches 565; Indels 30; Gaps 4;

63 ATGAGGATCTTTGTGGGCAACACACCTCTTCCCTCCCATGGAGGATCTCGTCCACCATGGCC 182
 1 ATGAGCTCTCTTGTACAGCAACACCAAAATTTGCTTTGATCTTTTCAAGAGATAGGC 60
 123 AAAGCAAGCCCAACCCCAACCTCTTCTCTCCCATGGAGGATCTCGTCCACCATGGCC 182
 61 AAGATGATCTGTATAAACAATATTTTCTCTCCCTGAGGCTCTCAGTGGCCCTTGGT 120
 183 ATGCTCTACATGGGCTCCAGGGGAGCAGCCAGACAGATGGCCAAAGTGTCTTCAATTT 242
 121 ATGCTACGCTTGGGTGTAGAAAGTGCACATGATGATGATGATGATGATGATGATGAT 180
 243 AATGAAGTGGGAGCAATGCACTTA-----CCCATGATCTCCAGAGAACTTTACCA 294
 181 AACGAATTTCCAGATGAAGCAAGAACTGACCTTGTCTGAAAGCAACAAACA 240
 295 GCTGTGGTTCATGACAGAGA--TCAGAAAGGATGATATCTGATGCGATTTTGCAGGCA 353
 241 AAAGTGTGGTGTACAGCTCTCTGGAGGGGAGCAAAAAAACACAGAGGCTCTGGATCAG 300
 354 CAAGCTGCAGATAAATCCAT-----TCATCTTCCGCTCTCTCAGC 395
 301 CAGCTGGTCTTAAACAATGAGACGAGCTGGTCACTGTCTTGGGAGCTTCTC 360
 396 TCTCAATCAATGATCCACAGGGAATTTTACTGGAAGTGTCAATTAAGCTTTGGT 455
 361 TCCAAATTAGACAGATCAAGACTGATTTACACACTGATGATTTGCCAACAGGCTTTATGA 420
 456 GAGAAGTCTCGAGCTTCGGGAGCAATATATTGCACTCTGTCAGAAATATTACTCTCA 515
 421 GAGCAGGAATTTCCCAATCTGTCAAGAACTTAGATGGTGTGATTCATTTTACCACAG 480
 516 GAACCCAGGAGTATGATCTCTAGAAATGTCAGAAAGCTAGAAAGAAAGATTAATTC 575
 481 ACGATTGAAGTGTGATTTCCAAAAAACCCCTGAAAAATCCAGCAAGAGATTAATTC 540
 576 TGGTCTCAGACTCAAAACCAAGGCAAAATCCCAACTTGTACCTGAAGTCTCTGTAGAT 635
 541 TGGTGTGATGTCAATCCCAAGTAAATCAAGGAATCTTTCAGCAAGGAGCTATTAAT 600
 636 GGGATACCAAGATGCTCTGTGATGATGCTGCTTCCGTTAACTCGCTCAGGCAACCTCA 695
 601 GCTGAGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 696 TTTGAGAGAACTAAATGGGCTTTATCTTCCGTTAACTCGCTCAGGCAACCTCA 755
 661 TTTGACCATGAACACGGTGTGATGACCTTTCTGCTAAATGCGATGAACAGAGT 720
 756 GTACAGATGATGTACTGCTGTAAGAGCTTAAACATTTGGATACATAGAGACCTTAAGGCT 815
 721 CTGAAGATGATGACCAAAAGGACTCTACAGAAATTTGGCTTCATAGAGGAGGTGAAGCA 780
 816 CAGATCTTAGAATCCCATATGCT---GGAGATGTTAGCATGTTCTTGTGCTTCCAGAT 872
 781 CAGATCTTGAATGAGGTACCAAGGGAGCTCAGCATGTTCTGCTGCTGCTGCTGCTGCT 840
 873 GAAATGCGGATGTGTCTACTGCTGAGAGCTCTGCTGAAAGTGAATTAACCTATGACAAA 932
 841 CACTTAAAGATAAAGCTGAAGGCTCTGGAAGAGCTTTGAAAGGAAATCACTATGAAAAA 900
 933 CTCACAAAGTGGACGACCAAGCAAAATGGCTGAAGATGAAGTGTGAGGTATACATACCC 992

Db 901 ATGTTGGCTGGAGCAGCTCAGAAAACATCTCAGAGAATCGTGTCTCTCTCTCCC 960
 Qy 993 CAGTTCAAATTAGAAGAGCAATATGAACTCAGATCCATCTCTGAAAGCATGGGATGAG 1052
 Db 961 CGGTTCAACCTGGAGACAGCTATGATCTCAATTCATTTTACAAGACATGGGCAATTACG 1020
 Qy 1053 GACGCTTTCAACAAGGGAGCGGCCAAATTTCTCAGGATGTGGAGAGAAATGACTGTTT 1112
 Db 1021 GATATCTTTGATGAACGAGGGCTGATCTTACTGGAATCTCTCCAAGTCCCAATTTGTAC 1080
 Qy 1113 CTTTCTGAAGTGTTCACCAAGCCATGGTGGATGTGAATGAGGAGGCACTGAAGCAGCC 1172
 Db 1081 TTGTCAAAATATATCCCAAAACCTTTGTGGAGGTGGATGAAAACCGTATCCCGAGCT 1140
 Qy 1173 GCTGCACAGAGGTGTATGACAGGGAATCTGAGACATGGAGGCCCAACAGTTTGTGSCA 1232
 Db 1141 GCAGCCACTGGGCTGTGTCTCGAAAGGTCACTACGATCTTTGGTGGAGTTTAAATGCC 1200
 Qy 1233 GATCATCGTTTCTTTTCTTATTTATGATGAATGAATGAATGAATGAATGAATGAATGA 1292
 Db 1201 AACCAACCTTTTCTCTTTTCAATTAAGACAAACCAAAACCAATCTCTCTTTTATGGC 1260
 Qy 1293 AGATTTTCTCACCCCTAA 1310
 Db 1261 AGGCTCTCTCTCTTAA 1278

RESULT 12
 US-08-315-831A-12
 ; Sequence 12, Application US/08315831A
 ; Patent No. 5663299
 ; GENERAL INFORMATION:
 ; APPLICANT: Remold-O'Donnell, Eileen
 ; TITLE OF INVENTION: Human Monocyte Elastase Inhibitor
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Wolf, Greenfield & Sacks, P.C.
 ; STREET: 600 Atlantic Ave.
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/315,831A
 ; FILING DATE: September 30, 1994
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: C0279/7016
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-720-3500
 ; TELEFAX: 617-720-2441
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1316 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 49..1188

OTHER INFORMATION: /codon_start= 49
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 49..1185
US-08-315-831A-12

Query Match 14.1%; Score 268.4; DB 1; Length 1316;
Best Local Similarity 55.3%; Pred. No. 1.5e-55;
Matches 543; Conservative 0; Mismatches 436; Indels 3; Gaps 1;

QY 341 GATTTTCAGGCAAGCTGCAGATAAAATCCATTTCCTCCGCTCTCTCAGCTCTGC 400
DB 216 GACTTTCCATTCAACACGGTGAAGAGGTTCAITTCAGATTCCAGAGTCTGAATGCTGA 275
QY 401 AATCAATGCATCCACAGGGAATTTTACTGGAAGTCTCAATAGCTGTTGGTGAGAA 460
DB 276 TATCAACAACAGTGGAGCGCTTATATCTGAAACTTGTCTAATAGATTATATGGAGAGAA 335
QY 461 GTCTCGAGCTTCGGGAAGAATATATTCGACTCTGTCAGAAATATTACTCTCTCAGAAC 520
DB 336 AACTTACAATTTCTCTCTGAGTTCTTGGTTTCGACTCAGAAACATATGGTGTGACCT 395
QY 521 CCAGGCAGTAGACTTCTAGAACTGTCAGAGAAGCTAGAAAAAGATTAAATTCCTGGGT 580
DB 396 GCGCAGTGTGATTTTCAGCATGCTCTGAGATGCAAGAGACCATTAACCACTGGGT 455
QY 581 CAAGACTCAACCAAGGCAAAATCCCAACTTGTACCTGAAGTTCTGTAGATGGGA 640
DB 456 CAAGGACACAGAGAGGAAAAATTCGGAACTGTGTGGCTTCGGGCATGGTTGATAACAT 515
QY 641 TACCAGATGGTCTGTGTAATGTCTACTCTCAAGAGAAAGTGGAAAACTCCATTTGA 700
DB 516 GACCAAACTGTGTAGTAAATGCCATCTATTCAAGGGAACCTGGAAGGATAAATTCAT 575
QY 701 GAAGAACTAAATGGGCTTTATCTCTCCGCTGTAACCTCGCTCAGCGCACACCTGTACA 760
DB 576 GAAGAAGCCACAGCAATGCAACATTCAGATTGAAATGAAGAAGACAAAACTGTGA 635
QY 761 GATGATGATCTCGGTGAAAGCTAAACATTGGATACATAGAAAGCTTAAAGGCTCAGAT 820
DB 636 AATGATGTATCAGAAGAAAAATTTGTCATATGGCTACATCGAGGACCTTAAAGTGGCGGT 695
QY 821 TCTAGAACTCCCATATCTCGAGATG---TTAGCATGTTCTTGTGCTTCAGATGAAT 877
DB 696 GCTGGAACTGCTTACCAAGCGAGGAGCTCAGCATGCTCTCTGCTCGCGATGACAT 755
QY 878 TGCCGATGTGCTACTGGCTTGGAGCTGCTGGAAGTGAATAACCTATGACAACTCAA 937
DB 756 TGAGGACGAGTCCACGGGCTTGAAGAAGATTGAGAACAGTTGACTTTGGAAAAGTTGCA 815
QY 938 CRAAGTGACCAAGCAAGCAAAATGGCTGGAAGATGAAGTTGAGGTATACATACCCAGTT 997
DB 816 TGAGTGGACTAAACCTCAGAACTCTCGATTTTCAATGAAGTTAATGTGCTGCTGCGGATG 875
QY 998 CAAATTAGAGAGCATTTAGAACTCAGATCCATCTGAAAGCATGGGCATGGAGGACGC 1057
DB 876 CAACTGGAAGAGATTACCTCTCACTCCGCTCGACCTCGCCGCTAGTGTGACGATCT 935
QY 1058 CTTCAACAAGGAGCGGCCAAATTTCTCAGGATGTGAGAGGAATGACCTGTTCTTTTC 1117
DB 936 CTTTAAACAGTAGAAGGCTGATCTGTCTGGCATGTCCAGAGCCAGAGATATTTTATATC 995
QY 1118 TGAAGTGTTCACCAAGCCATGTTGGATGTGAATGAGGAGGCACTGAAGCAGCGCGCTGG 1177
DB 996 AAAAAATTGTCACAAGTCAITTTGTGGAAGTGAATGAAGAGGGAACAGAGGCGGACGTCG 1055
QY 1178 CACAGAGGTGTTATGACGGAGAACTGACATGAGAGGCCCACTGTTGGCAGATCA 1237
DB 1056 CACAGAGGATCGCAACTTTCTGCATGTTGATGCCCAAGAAATTTCACTGCCGACCA 1115
QY 1238 TCCGTTTCTTTTCTTATATGCAATGAAGAACCAAGTGCATTTTATTTTTCGGGAGATT 1297
DB 1116 TCCATTCCTTTCTTATTCGGCATAATTCCTCAGTAGCATCTTATCTTCTTGGGAGATT 1175

QY 1298 TTGCTCACCCCTAAAACCTAAGCG 1319
DB 1176 TTCTTCCCTTAGAAGAAGAG 1197

RESULT 13
US-08-662-318-12
; Sequence 12, Application US/08662318
; Patent No. 5827672
; GENERAL INFORMATION:
; APPLICANT: Remold-O'Donnell, Eileen
; TITLE OF INVENTION: Human Monocyte Elastase Inhibitor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Ave.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/662,318
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/315,831
; FILING DATE: September 30, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: C0275/7016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1316 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 49..1188
; OTHER INFORMATION: /codon_start= 49
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 49..1185
; US-08-662-318-12

Query Match 14.1%; Score 268.4; DB 1; Length 1316;
Best Local Similarity 55.3%; Pred. No. 1.5e-55;
Matches 543; Conservative 0; Mismatches 436; Indels 3; Gaps 1;

QY 341 GATTTTCAGGCAAGCTGCAGATAAAATCCATTTCCTCCGCTCTCTCAGCTCTGC 400
DB 216 GACTTTCCATTCAACACGGTGAAGAGGTTCAITTCAGATTCCAGAGTCTGAATGCTGA 275
QY 401 AATCAATGCATCCACAGGGAATTTTACTGGAAGTCTCAATAGCTGTTGGTGAGAA 460
DB 276 TATCAACAACAGTGGAGCGCTTATATCTGAAACTTGTCTAATAGATTATATGGAGAGAA 335
QY 461 GTCTCGAGCTTCGGGAAGAATATATTCGACTCTGTCAGAAATATTACTCTCTCAGAAC 520

```

336 AACCTACAAATTCCTTCCTGAGTCTCTGGTTTCGACTCAGAAACATATGCTGCTGACCT 395
521 CCAGCAGTAGACTTCTTAGAATGTCAGAGAAGCTAGAAAAAGATTAATTCCTGGGT 580
396 GCGCAGTGTGATTTTCAGCATGCCCTCTGAAGATGCAAGAGACCAATAAACCAAGTGGT 455
581 CAAGACTCAAAACCAAGGCAAAATCCAACTGTTTACCTGTAAGAGTCTCTAGATGGGA 640
456 CAAGGACAGACAGAGGAAATTCGGAACCTGTTGGCTTCGGGCATGTTGATACAT 515
641 TACAGAGTGTCTGCTGTAATCTCTACTTCAAGAGGAAAGTGGAAAACTCCATTGA 700
516 GACCAAACTTGTGTAGTAATGCCATCTATTTCAGAGGAACTGGAGGATAAATTCAT 575
701 GAAGAACTTAATGGCTTTATCTCTTCCTGTTGTAAGTGAATGAAGAGACAGAAAACTGTAA 760
576 GAAGAGCCACGACGATGCACCATTCAGATTGAATGAAGAGACAGAAAACTGTAA 635
761 GATGATGATCTGCTGCGTAAAGCTAAACATATGATACATAGAGACCTTAAGGCTCAGAT 820
636 AATGATGATATCAGAGAAAAAATTTGCATATGCTACATCGAGGACCTTAAGTCCGCTGT 695
821 TCTAGAACTCCATATGCTGGAGATG--TTACGATGTTCTTGTGCTTCCAGATGAAT 877
696 GCTGGAATGCTTACCAAGGCGAGAGCTCAGATGCTCATCTGCTGCGGATGACAT 755
878 TGCCGATGTGCCACTGGCTTGGAGCTGTGGAAAGTGAATAACCTATGACAACTCAA 937
756 TGAGACGAGTCCACGGGCTGGAAGATTTGAGAACAGCTTGACTTTGGAAGTTGCA 815
938 CAAGTGGACCAAGACAAATGCTGGAAGATGAGTTGAGGTATACATACCCAGTT 997
816 TGAGTGGACTAACTGAGAACTCGAATTCATTTGAAATTAATGTCAGCTTGCCAGTT 875
998 CAATTAAGAGCATATTAAGTCTGATGCTTCTGAAAGCATGGCATGGAGGACGC 1057
876 CAACTGGAAGAGAGTTACACTCTCACTCCGACCTCGCCGCTAGGTGTCAGGATCT 935
1058 CTTCAACAGGACGCGCCAAATTTCTCAGGAGTGGAGAGGATGAGTATACATACCCAGTT 1117
936 CTTTAACTAGTAGAGGCTGATCTGCTGCAATGTCAGAGGCGAGATATTTTATATC 995
1118 TGAAGTGTCCACCAAGCATGTTGATGTAATGAGAGGCGCACTGAAGCAGCGCTGG 1177
996 AAAAAATTGCCAAGTCAATTTGTGGAAGTGAATGAAGGGAACAGAGGCGGAGCTGC 1055
1178 CACAGAGGTGTATCAGAGGAGAACTGCACATGAGGCGCCAGTTTGTGGCAGATCA 1237
1056 CACAGAGCATCGCACTTTCTGATGTTGATGCGCGAAGAAATTTCACTGCGGACCA 1115
1238 TCCGTTCTCTTTTCTTATATGATAGATAACCAAGTGCATTTTATTTTTCGCGAGAT 1297
1116 TCCATTCCTTTCTTTATTCGGCATAATTCCTCAGGTAGCATCTATCTTGGGGAGATT 1175
1298 TTGCTCACCTAAACTAAGC 1319
1176 TCTTCCCTTAGAAGAAAGAG 1197

```

RESULT 14

PCT-US95-12509-12

; Sequence 12, Application PC/TUS9512509

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Human Monocyte Elastase Inhibitor

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

; STREET: 600 Atlantic Ave.

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/12509
; FILING DATE: herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/315,831
; FILING DATE: 30 September 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: C0279/7016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1316 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 49..1188

; OTHER INFORMATION: /codon_start= 49

; FEATURE:

; NAME/KEY: mat_peptide

; LOCATION: 49..1185

; PCT-US95-12509-12

Query Match 14.1%; Score 268.4; DB 5; Length 1316;

Best Local Similarity 55.3%; Pred. No. 1.5e-55;

Matches 543; Conservative 0; Mismatches 436; Indels 3; Gaps 1;

```

QY 341 GATTTGAGGACACAGCTGCAGATAAAATCCATTCAATCCCTTCGCTCTCAGCTCTGC 400
DB 216 GACTTTCATTTCAACACAGGTTCAAGAGGTTCAATCAAGATTCAGAGTCTGAATGCTGA 275
QY 401 ATCAATGCATCCACAGGGAATTTATTACTGGAAGTGTCAATAAGCTGTTTGGTGAGAA 460
DB 276 TATCAACAAACGTGGAGCGTCTTATATCTGAAACTTGTCTAATAGATTATATGGAGAGAA 335
QY 461 GTCTCGAGCTTCGCGGAAGAATATATTCGACTCTGTCAAGAAATATTACTCCTCAGAAC 520
DB 336 AACTTACAATTTCTTCTGAGTTCTTGGTTTCGACTCAGAAAAACATATGTTGCTGACCT 395
QY 521 CCAGGAGTAGACTTCCCTAGAAATGTCAGAGAAGCTAGAAAAAGATTAAATTCCTGGGT 580
DB 396 GCGCAGTGTGATTTTCAGCATGCCCTCTGAAGATGCAAGAGACCAATAAACCAAGTGGT 455
QY 581 CAAGACTCAAAACCAAGGCAAAATCCCAAATCTGTTTACCTGAAAGTCTCTAGATGGGA 640
DB 456 CAAAGGACAGACAGAGGAAAAATTCGGAACCTGTTGGCTTCGGGCATGTTGATAACAT 515
QY 641 TACCGAGTGTCTGCTGTAATGCTGTCTACTTCAAGAGGAAAGTGGAAAACTCCATTGA 700
DB 516 GACCAAACTTGTGTAGTAATGCCATCTATTTCAAGGGAACCTGGAAGGATAAATTCAT 575
QY 701 GAAGAACTTAATGGGCTTTATCTCTTCCGTGTAACCTCGCTCAGGCGCACACCTGTACA 760
DB 576 GAAAGAGCCACGACGAAATGCACCAATTCAGATTGAATGAAGAGACAGAAAACTGTGAA 635
QY 761 GATGATGATCTTCGCTGGAAGAGCTTAACATTTGATACATAGAACCTTAAGGCTCAGAT 820

```

Db 636 AATGATGATATCAGAGAAAAAATTTGCATATGCTACATCGAGGACCTTAAGTCGCGTGT 695
Qy 821 TCTAGAACTCCCATATGCTGGAGATG---TTAGCATGTTCTTGTGCTTCCAGATGAAT 877
Db 696 GCTGGAACATGCTTTACCAAGCGGAGGAGCTCAGCATGCTGCTGCTGCGGATGACAT 755
Qy 878 TCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 937
Db 756 TGAGGACGAGTCCACGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 815
Qy 938 CAAGTGGACCGCAAGACCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
Db 816 TGAGTGGATTAACCTGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 875
Qy 998 CAAATTTAGAGACATTTAGATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1057
Db 876 CAAATTTAGAGACATTTAGATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 935
Qy 1058 CTTTCAACAGGAGGCGGCGCAATTTCTCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1117
Db 936 CTTTAAACAGTAGCAAGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 995
Qy 1118 TGAAGTGTTCACCAAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177
Db 996 AAAAATTTGCTCAACAGTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1055
Qy 1178 CACAGGAGGTGTTATGACAGGAGAACTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1237
Db 1056 CACAGCAGCATCGCAACTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1115
Qy 1238 TCGGTTCTTTCTTATTTATGATAGATAAACAAGTGCATTTTATTTTTCGCGCAGATT 1297
Db 1116 TCAATTCCTTTCTTATTCGCGATATTTCTCAGGTAGCATCTTATTTCTTGGGAGATT 1175

RESULT 15

US-09-016-434-710
; Sequence 710, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Selhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166 710:
; INFORMATION FOR SEQ ID NO:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1430 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: EOSIHET02
; CLONE: 334703
; US-09-016-434-710

Query Match 14.1%; Score 268.4; DB 4; Length 1430;
Best Local Similarity 55.3%; Pred. No. 1.6e-55;
Matches 543; Conservative 0; Mismatches 436; Indels 3; Gaps 1;
Qy 341 GATTTTTCGAGGCAACAGCTGCAGATAAAATCCATTCCTTCCGCTCTCTCAGCTCTGC 400
Db 348 GACTTTCCATTTCAACACGGTTGAAGAGGTTCAATCAAGATTCCAGAGTCTGAATGCTGA 407
Qy 401 AATCAATGATCCACAGGGNAATTTTACTGGAAGTGTCAATAGCTGTTTGGTGAGAA 460
Db 408 TATCAACAAACGTGGAGCGCTTATATCTGAAACTTGTCTAATAGATTATATGGAGAGAA 467
Qy 461 GTCTGCGAGTTCGCGGAGAGATATATTCGACTCTGTCTGAGAAATATTACTCTCAGAAC 520
Db 468 AACTTCAAAATTTCTTCTTCTGAGTTCTTGGTTTCGACTCAGAAAAACATATGTTGCTGAC 527
Qy 521 CCAGGCGAGTAGACTTCTCTAGAAATGTGAGAGAGAGCTAGAAAAAGATTAATTCCTGGGT 580
Db 528 GGCAGTGTGGATTTTTCAGCATGCTCTGAGATGCAAGGAGACCATTAACAGTGGT 587
Qy 581 CAAGACTCAAAACAAAGGCAAAATCCCAAACCTTTTACCTGAAAGTTCCTGAGATGGGA 640
Db 588 CAAAGGACAGACAG 647
Qy 641 TACCGAGTGTCTGCTGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
Db 648 GACCAAACTGTGCTAGTAATGCCATCTATTTCAAGGGAACCTGGAAGGATAAATTCAT 707
Qy 701 GAAGAACTAAATGGGCTTTTATCTTCCGTTGTAATCTCGCTCAGCCACACCTGTACA 760
Db 708 GAAAGAGCCACACGAAATCCCAATTCAGATTGAATAAGAAAGACAGAAAACTGTGAA 767
Qy 761 GATGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 820
Db 768 AATGATGATCAAG 827
Qy 821 TCTAGAACTCCCATATGCTGGAGATG---TTAGCATGTTCTTGTGCTTCCAGATGAAT 877
Db 828 GCTGGAACATGCTTACCAAGCGGAGGAGCTCAGCATGCTGCTGCTGCTGCTGCTGCTGCT 887
Qy 878 TGCCGATGTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 937
Db 888 TGAGGACGAGTCCACGGGCTGGAAGAGATTGAGGAAACAGTTGACTTTTGGAAAAAGTTGCA 947
Qy 938 CAAAGTGGACCAAGACAGCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
Db 948 TGAGTGGATTAACCTGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1007
Qy 998 CAAATTTAGAGACATTTAGATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1057
Db 1008 CAAACTGGAAGAGAGATTACATCTCTCAACTCCGCTCCCGCCCTAGGTGTGAGATCT 1067
Qy 1058 CTTTCAACAGGAGGCGGCGCAATTTCTCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1117
Db 1068 CTTTAAACAGTAGCAAGGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1127
Qy 1118 TGAAGTGTTCACCAAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177
Db 1128 AAAAATTTGCTCAACAGTCAATTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1187

```

2y 1178 CACAGGAGGTGTATGACAGGGAGAACTGGACATGGAGGCCCCACAGTTTGTGCGAGATCA 1237
    |||||
2b 1188 CACAGCAGGCATCGCACTTCTGCAATGTTGATGCCCGAAGAAAATTCACCTGCCGACCA 1247
    |||||
2y 1238 TCCGTTTCTTTTCTTATTATGCATAAGATAACCAAGTGCATTTTATTTTTCGGCGAGATT 1297
    |||||
2b 1248 TCCATTCTTTTCTTTTATTCGGCATAATTCCTCAGGTAGCATCCTATTCTTGGGGAGATT 1307
    |||||
2y 1298 TTGCTCACCCCTAAACTAAGCG 1319
    |||||
2b 1308 TTCTTCCCTTAGAAGAAAGAG 1329
    |||||

```

Search completed: March 11, 2004, 14:41:44
Job time : 93.1303 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model

un on: March 11, 2004, 03:37:21 ; Search time 444.601 Seconds
(without alignments)
15799.545 Million cell updates/sec

itle: US-10-084-817-27

erfect score: 1908
equence: 1 gagagcattgccgcgcaga.....aacaaaaataaaaaaagg 1908

oring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

earched: 2432557 seqs, 1840798884 residues

otal number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	1908	100.0	1908	14	US-10-084-817-27
2	1897	99.4	1897	14	US-10-208-408-17
3	1896	99.4	1935	14	US-10-106-698-261
4	1872.6	98.1	1900	12	US-10-240-425-1566
5	985.8	51.7	191395	12	US-10-235-192A-45
6	984.8	51.6	1199	9	US-09-962-832-252
7	560	29.4	566	9	US-09-813-358-178
8	560	29.4	566	9	US-09-997-279-178
9	498.8	26.1	503	9	US-09-813-358-66
10	498.8	26.1	503	10	US-09-997-279-66
11	475.6	24.9	482	9	US-09-864-761-13940
12	387.6	20.3	596	14	US-10-029-386-13440
13	350.2	18.4	1256	15	US-10-027-632-122703
14	343.2	18.0	354	14	US-10-029-386-27140
15	337.2	17.7	544	14	US-10-029-386-865

16	317.8	16.7	1275	14	US-10-113-113-3	Sequence 3, Appl 1
17	315.4	16.5	317	9	US-09-864-761-30504	Sequence 30504, A
18	315.4	16.5	317	14	US-10-029-386-14568	Sequence 14568, A
19	276.8	14.5	3603	14	US-10-113-113-1	Sequence 1, Appl 1
20	274.4	14.4	314	9	US-09-813-358-186	Sequence 186, Appl 1
21	274.4	14.4	314	10	US-09-997-279-186	Sequence 186, Appl 1
22	272.4	14.3	1278	14	US-10-094-944-1	Sequence 1, Appl 1
23	271	14.2	1200	10	US-09-823-187-1	Sequence 1, Appl 1
24	269	14.1	1185	12	US-10-258-951-20	Sequence 20, Appl 1
25	268.4	14.1	1316	9	US-09-969-347-192	Sequence 192, Appl 1
26	268.4	14.1	1316	12	US-10-240-425-1339	Sequence 1339, Ap
27	268.4	14.1	1430	15	US-10-305-720-710	Sequence 710, App
28	268.4	14.1	1458	14	US-10-101-510-347	Sequence 347, App
29	268.4	14.1	1931	9	US-09-925-301-358	Sequence 358, App
30	259.2	13.6	2221	14	US-10-198-846-13963	Sequence 13963, A
31	252.4	13.2	1385	15	US-10-305-720-1075	Sequence 1075, Ap
32	252.4	13.2	1624	10	US-09-993-363-1	Sequence 1, Appl 1
33	250.8	13.1	1624	9	US-09-925-300-500	Sequence 500, App
34	250.8	13.1	1707	14	US-10-084-817-148	Sequence 148, App
35	249.2	13.1	1361	15	US-10-305-720-1284	Sequence 1284, Ap
36	249.2	13.1	1465	9	US-09-880-107-3027	Sequence 3027, Ap
37	240	12.6	1173	15	US-10-094-886-15	Sequence 15, Appl 1
38	237.4	12.4	241	9	US-09-604-287A-338	Sequence 338, App
39	237.4	12.4	241	10	US-09-551-621-338	Sequence 338, App
40	237.4	12.4	241	13	US-10-007-805-338	Sequence 338, App
41	237.4	12.4	241	14	US-10-076-622-338	Sequence 338, App
42	237.4	12.4	241	14	US-10-124-805-338	Sequence 338, App
43	234.8	12.3	1279	14	US-10-157-031-354	Sequence 354, App
44	234.8	12.3	1308	9	US-09-735-705-109	Sequence 109, App
45	234.8	12.3	1308	9	US-09-850-716A-109	Sequence 109, App

ALIGNMENTS

RESULT 1

US-10-084-817-27
; Sequence 27, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Ued G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; PRIOR FILING DATE: 2002-02-25
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 27
; LENGTH: 1908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 063646CB1
US-10-084-817-27

Query Match 100.0%; Score 1908; DB 14; Length 1908;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1908; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGAGCATTGCCCGTCAGACGACAACTCAGAGAATAACAGAGAACACAGATTGAAA 60
DB 1 GAGGAGCATTGCCCGTCAGACGACAACTCAGAGAATAACAGAGAACACAGATTGAAA 60
QY 61 CAATGGAGGATCTTTGTGGCAACACACTTTTCCCTCAATTTTCAAGCATCTGG 120
DB 61 CAATGGAGGATCTTTGTGGCAACACACTTTTCCCTCAATTTTCAAGCATCTGG 120

		Matches 1897; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
2y	1	GAGGAGCAATGCCCCGTCAGACAGCACTCAGAGAAATACAGAGAAACACAGATTGAAA	60		
2b	1	GAGGAGCAATGCCCCGTCAGACAGCACTCAGAGAAATACAGAGAAACACAGATTGAAA	60		
2y	61	CAATGGAGGATCTTTGTGTGGGAAAACACACTCTTTTGCCTCAATTTATTTCAAGCATCTGG	120		
2b	61	CAATGGAGGATCTTTGTGTGGGAAAACACACTCTTTTGCCTCAATTTATTTCAAGCATCTGG	120		
2y	121	CAAAAGCAAGCCCAACCAAGCACTTCTCTCCCATGGAGCAATCTCTCCACCATGG	180		
2b	121	CAAAAGCAAGCCCAACCAAGCACTTCTCTCCCATGGAGCAATCTCTCCACCATGG	180		
2y	181	CCATGGTCTACATGGGCTCCAGGGGAGCAACGAGACCCAGAGTGGCCAAAGTGTTCAGT	240		
2b	181	CCATGGTCTACATGGGCTCCAGGGGAGCAACGAGACCCAGAGTGGCCAAAGTGTTCAGT	240		
2y	241	TTAATGAAGTGGAGCCCAATGAGTTACCCCATGATCTCCAGAGAACTTTACCAAGCTGTG	300		
2b	241	TTAATGAAGTGGAGCCCAATGAGTTACCCCATGATCTCCAGAGAACTTTACCAAGCTGTG	300		
2y	301	GGTTCATCAGCAGATCCAGAGAGGTAGTTATCCTGATGCGATTTTCAGGCAAAAGCTG	360		
2b	301	GGTTCATCAGCAGATCCAGAGAGGTAGTTATCCTGATGCGATTTTCAGGCAAAAGCTG	360		
2y	361	CAGATAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGA	420		
2b	361	CAGATAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGA	420		
2y	421	ATTATTTACTGGAAGTGTCAATAAGCTGTTGGTGAGAAAGTCTGCGAGCTTCGGGAAG	480		
2b	421	ATTATTTACTGGAAGTGTCAATAAGCTGTTGGTGAGAAAGTCTGCGAGCTTCGGGAAG	480		
2y	481	AATATATTCGACTCTGTGAGAAATATTAATCTCTCAGAACCCAGGAGTACCTCTAG	540		
2b	481	AATATATTCGACTCTGTGAGAAATATTAATCTCTCAGAACCCAGGAGTACCTCTAG	540		
2y	541	AATGTGCAAGAGCTAGAAAAGATTAATTCCTGGTCAAGACTCAAAACCAAGGCA	600		
2b	541	AATGTGCAAGAGCTAGAAAAGATTAATTCCTGGTCAAGACTCAAAACCAAGGCA	600		
2y	601	AAATCCCAAACTTGTTACCTGAGGTTCTGTAGATGGGATACAGAGTGGTCTGGTGA	660		
2b	601	AAATCCCAAACTTGTTACCTGAGGTTCTGTAGATGGGATACAGAGTGGTCTGGTGA	660		
2y	661	ATGCTGTCTACTTCAAGAGAGTGAAACTCCATTTGAGAGAACTAAATGGGCTTT	720		
2b	661	ATGCTGTCTACTTCAAGAGAGTGAAACTCCATTTGAGAGAACTAAATGGGCTTT	720		
2y	721	ATCCTTTCCGTGTAACCTCGGCTCAGGCACACCTGTACAGATGATGTACTTGGTGAAA	780		
2b	721	ATCCTTTCCGTGTAACCTCGGCTCAGGCACACCTGTGTACAGATGATGTACTTGGTGAAA	780		
2y	781	AGCTAAACATTCGATACATAGAGACTTAAGGCTCAGATTCAGACTCCCATATGCTG	840		
2b	781	AGCTAAACATTCGATACATAGAGACTTAAGGCTCAGATTCAGACTCCCATATGCTG	840		
2y	841	GAGATGTTAGCATGTTCTTTGCTTCCAGATGAAATTTGCGATGTGTCCACTGGCTGG	900		
2b	841	GAGATGTTAGCATGTTCTTTGCTTCCAGATGAAATTTGCGATGTGTCCACTGGCTGG	900		
2y	901	AGCTCTCGAAGTGAATTAACCTATGACAACTCAACAAAGTGACCAAGCAACAA	960		
2b	901	AGCTCTCGAAGTGAATTAACCTATGACAACTCAACAAAGTGACCAAGCAACAA	960		
2y	961	TGGCTGAAGATGAAGTTGAGGTATACATACCCAGTTTCAATTTAGAGAGCATTTAGAC	1020		
2b	961	TGGCTGAAGATGAAGTTGAGGTATACATACCCAGTTTCAATTTAGAGAGCATTTAGAC	1020		
2y	1021	TCAGATCCATTTCTGAAAAGCATGGGATGAGAGCGCTTCAACAAAGGAGCGGCCAATT	1080		
2b	1021	TCAGATCCATTTCTGAAAAGCATGGGATGAGAGCGCTTCAACAAAGGAGCGGCCAATT	1080		

RESULT 3

US-10-106-698-261
; Sequence 261, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005PI
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280

Qy	1081	TCCTAGGATGTCGAGAGGAATGACCTGTTCTTTCTGAGTGTTCACCAAGCCATGG	1140
Db	1081	TCCTAGGATGTCGAGAGGAATGACCTGTTCTTTCTGAGTGTTCACCAAGCCATGG	1140
Qy	1141	TGGATGTGAATGAGGAGGCACTGAAGCAGCCGCTGGCACAGGAGGTGTTATGACAGGA	1200
Db	1141	TGGATGTGAATGAGGAGGCACTGAAGCAGCCGCTGGCACAGGAGGTGTTATGACAGGA	1200
Qy	1201	GAACTGGACATGAGGAGCCACACAGTTTGTGCGAGATCATCCGTTCTTTTCTTATTATGC	1260
Db	1201	GAACTGGACATGAGGAGCCACACAGTTTGTGCGAGATCATCCGTTCTTTTCTTATTATGC	1260
Qy	1261	ATAAGATAACCAAGTGCATTTTATTTTCCGCGAGATTTGCTCACCTTAAACTAAGCGT	1320
Db	1261	ATAAGATAACCAAGTGCATTTTATTTTCCGCGAGATTTGCTCACCTTAAACTAAGCGT	1320
Qy	1321	GCTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCGCTCAGAAATGCTATTTCAAT	1380
Db	1321	GCTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCGCTCAGAAATGCTATTTCAAT	1380
Qy	1381	TGCCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAACTTCTGCTACC	1440
Db	1381	TGCCAAAATTTAGAGATGTTTCTACATATTTCTGCTCTTCTGAAACAACTTCTGCTACC	1440
Qy	1441	CACATAAATAAACAACAGAAATTAATAGACAAATGCTCTATTATTAACATGACAAACCTATT	1500
Db	1441	CACATAAATAAACAACAGAAATTAATAGACAAATGCTCTATTATTAACATGACAAACCTATT	1500
Qy	1501	AATCATTTGGTCTCTTAAATGGGATCATGCCATTTAGATTTTCTTACTATGAGTTTA	1560
Db	1501	AATCATTTGGTCTCTTAAATGGGATCATGCCATTTAGATTTTCTTACTATGAGTTTA	1560
Qy	1561	TTTTTATTAACATTAACCTTTTACTTTGTTATTTATTTATTTATTAATGAGTTTAA	1620
Db	1561	TTTTTATTAACATTAACCTTTTACTTTGTTATTTATTTATTTATTAATGAGTTTAA	1620
Qy	1621	ATTATGCTCAGTCCCTATTAATGAGTAAATAGTTATAGAGAGAGATGATCTGTTA	1680
Db	1621	ATTATGCTCAGTCCCTATTAATGAGTAAATAGTTATAGAGAGAGATGATCTGTTA	1680
Qy	1681	ATTTCTCTATTAATAAATGCTTTTAAATGTTCTCATTAATGAAGTAAGTAGTATCCCT	1740
Db	1681	ATTTCTCTATTAATAAATGCTTTTAAATGTTCTCATTAATGAAGTAAGTAGTATCCCT	1740
Qy	1741	CCATGCCCTTCTGTAATAAATATCTGAAAAAACATTAACAAATAGGCAATATATGTTA	1800
Db	1741	CCATGCCCTTCTGTAATAAATATCTGAAAAAACATTAACAAATAGGCAATATATGTTA	1800
Qy	1801	TGTGATTTCTAGAAATACATAACATATATATGCTGTATCTTATTCATTTGCAAG	1860
Db	1801	TGTGATTTCTAGAAATACATAACATATATATGCTGTATCTTATTCATTTGCAAG	1860
Qy	1861	TATATAATAAATAAACCTGCTTCCAAACAAACAAAAA	1897
Db	1861	TATATAATAAATAAACCTGCTTCCAAACAAACAAAAA	1897

PRIOR FILING DATE: 1999-11-03 NUMBER OF SEQ ID NOS: 8564 SOFTWARE: PatentIn Ver. 3.0 SEQ ID NO 261 LENGTH: 1935 TYPE: DNA ORGANISM: Homo sapiens US-10-106-698-261									
Query Match 99.4%; Score 1896; DB 14; Length 1935; Best Local Similarity 99.7%; Pred. No. 0; Matches 1899; Conservative 0; Mismatches 5; Indels 0; Gaps 0;									
QY	3	GGAGCAATGCCCTCAGACAGCAACTCAGAGAAATACAGAGAAACACAGATTGAACA	62	903	CTGCTGGAAGTGAATTAACCTATGACAACTCAACAACTGAGCCAGCAAAAGACAAATG	962	QY	903	CTGCTGGAAGTGAATTAACCTATGACAACTCAACAACTGAGCCAGCAAAAGACAAATG
DB	15	GGAGCAATGCCCTCAGACAGCAACTCAGAGAAATACAGAGAAACACAGATTGAACA	74	915	CTGCTGGAAGTGAATTAACCTATGACAACTCAACAACTGAGCCAGCAAAAGACAAATG	974	DB	915	CTGCTGGAAGTGAATTAACCTATGACAACTCAACAACTGAGCCAGCAAAAGACAAATG
QY	63	ATGGAGGATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCA	122	963	GCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAAATTAGAAAGAGCAATTATGAATC	1022	QY	963	GCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAAATTAGAAAGAGCAATTATGAATC
DB	75	ATGGAGGATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCA	134	975	GCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAAATTAGAAAGAGCAATTATGAATC	1034	DB	975	GCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAAATTAGAAAGAGCAATTATGAATC
QY	123	AAAGCAAGCCCAACCAAGACCTCTCTCCCATGGAGCATCTCGTCCACCATGGCC	182	1023	AGATCCATTCTGAAAAGCATGGGATGGAGAGCGCTTCAACAAGGGAGCGGCCCAATTTTC	1082	QY	1023	AGATCCATTCTGAAAAGCATGGGATGGAGAGCGCTTCAACAAGGGAGCGGCCCAATTTTC
DB	135	AAAGCAAGCCCAACCAAGACCTCTCTCCCATGGAGCATCTCGTCCACCATGGCC	194	1035	AGATCCATTCTGAGAAGCATGGGATGGAGAGCGCTTCAACAAGGGAGCGGCCCAATTTTC	1094	DB	1035	AGATCCATTCTGAGAAGCATGGGATGGAGAGCGCTTCAACAAGGGAGCGGCCCAATTTTC
QY	183	ATGGCTACATGGCTCCAGGGCAGACCCGAAAGACAGATGGCAAGGTGCTTCAGTTT	242	1083	TCAGGATGTCGAGAGGAAATGACCTGTTCTTTCTGAGTGTTCACCAAGCCATGGTG	1142	QY	1083	TCAGGATGTCGAGAGGAAATGACCTGTTCTTTCTGAGTGTTCACCAAGCCATGGTG
DB	195	ATGGCTACATGGCTCCAGGGCAGACCCGAAAGACAGATGGCAAGGTGCTTCAGTTT	254	1095	TCAGGATGTCGAGAGGAAATGACCTGTTCTTTCTGAGTGTTCACCAAGCCATGGTG	1154	DB	1095	TCAGGATGTCGAGAGGAAATGACCTGTTCTTTCTGAGTGTTCACCAAGCCATGGTG
QY	243	AATGAAGTGGAGCCCAATGAGTTACCCCATGACTCCAGAGAACTTTACAGAGCTGTGGG	302	1143	GATGTGAATGAGAGGGCACTGAAGAGCGCTCGCAGAGAGGTGTATGACAGGGAGA	1202	QY	1143	GATGTGAATGAGAGGGCACTGAAGAGCGCTCGCAGAGAGGTGTATGACAGGGAGA
DB	255	AATGAAGTGGAGCCCAATGAGTTACCCCATGACTCCAGAGAACTTTACAGAGCTGTGGG	314	1155	GATGTGAATGAGAGGGCACTGAAGAGCGCTCGCAGAGAGGTGTATGACAGGGAGA	1214	DB	1155	GATGTGAATGAGAGGGCACTGAAGAGCGCTCGCAGAGAGGTGTATGACAGGGAGA
QY	303	TTTCATCAGAGATCCAGAGGGTAGTTATCTGATGGGATTTTCAGGCGCAAGCTGCA	362	1203	ACTGGACATGGAGGCCACACAGTTTGTGGCAGATCATCGTTTCTTTCTTATTTATGCAAT	1262	QY	1203	ACTGGACATGGAGGCCACACAGTTTGTGGCAGATCATCGTTTCTTTCTTATTTATGCAAT
DB	315	TTTCATCAGAGATCCAGAGGGTAGTTATCTGATGGGATTTTCAGGCGCAAGCTGCA	374	1215	ACTGGACATGGAGGCCACACAGTTTGTGGCAGATCATCGTTTCTTTCTTATTTATGCAAT	1274	DB	1215	ACTGGACATGGAGGCCACACAGTTTGTGGCAGATCATCGTTTCTTTCTTATTTATGCAAT
QY	363	GATAAATCAATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAAT	422	1263	AGATTAACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCCTAAACCTAAGCGTGC	1322	QY	1263	AGATTAACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCCTAAACCTAAGCGTGC
DB	375	GATAAATCAATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAAT	434	1275	AGATTAACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCCTAAACCTAAGCGTGC	1334	DB	1275	AGATTAACCAAGTGCATTTATTTTTCGGCAGATTTTCTCACCCTAAACCTAAGCGTGC
QY	423	TATTTACTGGAAGTGCATTAAGCTTTTGGTGAGAGTCTGGAGCTTCGGGAAGAA	482	1323	TGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTTCAAAATG	1382	QY	1323	TGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTTCAAAATG
DB	435	TATTTACTGGAAGTGCATTAAGCTTTTGGTGAGAGTCTGGAGCTTCGGGAAGAA	494	1335	TGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTTCAAAATG	1394	DB	1335	TGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTTCAAAATG
QY	483	TATTTCTGACTCTGTGAGAAATATTTACTCTCAGAACCCCAAGGAGTAGACTTCTCAGAA	542	1383	CGAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTCTGGAACAATCTTCTGCTACCCA	1442	QY	1383	CGAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTCTGGAACAATCTTCTGCTACCCA
DB	495	TATTTCTGACTCTGTGAGAAATATTTACTCTCAGAACCCCAAGGAGTAGACTTCTCAGAA	554	1395	CGAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTCTGGAACAATCTTCTGCTACCCA	1454	DB	1395	CGAAAAATTTAGAGATGTTTCTACATATTTCTGCTCTCTGGAACAATCTTCTGCTACCCA
QY	543	TGTGCAAGAGCTAGAAAAGATTAATTCCTGSGTCAAGACTCAAAACCAAGGCAAA	602	1443	CTAAATTAACCAACAGAGAAATTAATTTAGCAATTTGTCTATTAATAACAGCAACCCCTATTAA	1502	QY	1443	CTAAATTAACCAACAGAGAAATTAATTTAGCAATTTGTCTATTAATAACAGCAACCCCTATTAA
DB	555	TGTGCAAGAGCTAGAAAAGATTAATTCCTGSGTCAAGACTCAAAACCAAGGCAAA	614	1455	CTAAATTAACCAACAGAGAAATTAATTTAGCAATTTGTCTATTAATAACAGCAACCCCTATTAA	1514	DB	1455	CTAAATTAACCAACAGAGAAATTAATTTAGCAATTTGTCTATTAATAACAGCAACCCCTATTAA
QY	603	ATCCCAACTTGTACTGAGAGTCTGTAGATGGGATACAGAGATGGTCTGTGTGAAT	662	1503	TCATTTGGTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATT	1562	QY	1503	TCATTTGGTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATT
DB	615	ATCCCAACTTGTACTGAGAGTCTGTAGATGGGATACAGAGATGGTCTGTGTGAAT	674	1515	TCATTTGGTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATT	1574	DB	1515	TCATTTGGTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTATT
QY	663	GCTGTCTACTTCAAGAGAGTGGAAACTCCATTTGAGAGAACTAAATGGGCTTTAT	722	1563	TTTATAACATTAACCTTTTACTTTTATTTATTTATTTATTAATAATGTTGAGTTTAAAT	1622	QY	1563	TTTATAACATTAACCTTTTACTTTTATTTATTTATTTATTAATAATGTTGAGTTTAAAT
DB	675	GCTGTCTACTTCAAGAGAGTGGAAACTCCATTTGAGAGAACTAAATGGGCTTTAT	734	1575	TTTATAACATTAACCTTTTACTTTTATTTATTTATTTATTAATAATGTTGAGTTTAAAT	1634	DB	1575	TTTATAACATTAACCTTTTACTTTTATTTATTTATTTATTAATAATGTTGAGTTTAAAT
QY	723	CCTTTCCGTGTAACTCCGGCTCAGGCACTGTGATGATGATGATGATGATGATGATGATGAT	782	1623	TATTTGCTCACTGCTATTTTAAATGAGTAAATAAGTTATAGAGCAGATGATCTGTTAAT	1682	QY	1623	TATTTGCTCACTGCTATTTTAAATGAGTAAATAAGTTATAGAGCAGATGATCTGTTAAT
DB	735	CCTTTCCGTGTAACTCCGGCTCAGGCACTGTGATGATGATGATGATGATGATGATGATGAT	794	1635	TATTTGCTCACTGCTATTTTAAATGAGTAAATAAGTTATAGAGCAGATGATCTGTTAAT	1694	DB	1635	TATTTGCTCACTGCTATTTTAAATGAGTAAATAAGTTATAGAGCAGATGATCTGTTAAT
QY	783	CTAAACATTTGGATACATAGAGACTTAAAGGCTCAGATTTCTAGAACTCCCATATGCTGGA	842	1683	TTCCCTATCTAAATAATGCTTTTAAATGCTTCTCATAAATGAGAAATAGTAGGATCCCTCC	1742	QY	1683	TTCCCTATCTAAATAATGCTTTTAAATGCTTCTCATAAATGAGAAATAGTAGGATCCCTCC
DB	795	CTAAACATTTGGATACATAGAGACTTAAAGGCTCAGATTTCTAGAACTCCCATATGCTGGA	854	1695	TTCCCTATCTAAATAATGCTTTTAAATGCTTCTCATAAATGAGAAATAGTAGGATCCCTCC	1754	DB	1695	TTCCCTATCTAAATAATGCTTTTAAATGCTTCTCATAAATGAGAAATAGTAGGATCCCTCC
QY	843	GATGTAGCATGTTCTTTGTTGCTTCCAGATGAAATGCGGATGTTCCACCTGGCTTGAG	902	1743	ATCCCTTCTGTAAATAATCTCGAAAAAATCAATTAACCAATAGGCAATATATGTTATG	1802	QY	1743	ATCCCTTCTGTAAATAATCTCGAAAAAATCAATTAACCAATAGGCAATATATGTTATG
DB	855	GATGTAGCATGTTCTTTGTTGCTTCCAGATGAAATGCGGATGTTCCACCTGGCTTGAG	914	1755	ATCCCTTCTGTAAATAATCTCGAAAAAATCAATTAACCAATAGGCAATATATGTTATG	1814	DB	1755	ATCCCTTCTGTAAATAATCTCGAAAAAATCAATTAACCAATAGGCAATATATGTTATG
1803 TGCATTTCTAGAAATACATTAACATATATATGCTGATCTTATATTAATCAATGCAAGTA 1862 1815 TGCATTTCTAGAAATACATTAACATATATATGCTGATCTTATATTAATCAATGCAAGTA 1874 1863 TATAAATAAATCACTGCTTCCAAACCAACAAAAATAAAAAA 1906 1875 TATAAATAAATCACTGCTTCCAAACCAACAAAAATAAAAAA 1918									

RESULT 4
 US-10-240-425-1566
 ; Sequence 1566, Application US/10240425
 ; Publication No. US20040033502A1

GENERAL INFORMATION:

APPLICANT: Williams, Amanda
 APPLICANT: Boland, Joseph F.
 APPLICANT: Lord, Reginald V.
 APPLICANT: Alvarez, Chris
 APPLICANT: Wetzel, Jon C.
 APPLICANT: Scherf, Uwe
 APPLICANT: Vockley, Joseph G.
 TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
 FILE REFERENCE: 44921-5026
 CURRENT APPLICATION NUMBER: US/10/240,425
 CURRENT FILING DATE: 2002-09-30
 PRIOR APPLICATION NUMBER: PCI/US01/09847
 PRIOR FILING DATE: 2001-03-28
 PRIOR APPLICATION NUMBER: US 60/193,446
 PRIOR FILING DATE: 2000-03-31
 NUMBER OF SEQ ID NOS: 1588
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1566
 LENGTH: 1900

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Genbank Accession No. US20040033502A1 Y00630
 IS-10-240-425-1566

Query Match 98.1%; Score 1872.6; DB 12; Length 1900;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1886; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

2y	1	GAGGAGCATTCGCGTCAGACAGCACTCAGAGATAACACAGAGAACACACAGATTGAAA	60
2b	11	GAGGAGCATTCGCGTCAGACAGCACTCAGAGATAACACAGAGAACACACAGATTGAAA	70
2y	61	CAATGGAGGATCTTTGTTGTCGCAACACACTCTTTGCCCTCAATTTATTCAGACATCTGG	120
2b	71	CAATGGAGGATCTTTGTTGTCGCAACACACTCTTTGCCCTCAATTTATTCAGACATCTGG	130
2y	121	CAGAAGCAAGCCCAACAGCACTCTTCTCTCCCATGGAGCATCTCGTCCACCATGG	180
2b	131	CAGAAGCAAGCCCAACAGCACTCTTCTCTCCCATGGAGCATCTCGTCCACCATGG	190
2y	181	CCATGGTCTACATGGGTCCTCAGGGGAGCAGCAGAGACAGATGGCCAGGTCCTCAGT	240
2b	191	CCATGGTCTACATGGGTCCTCAGGGGAGCAGCAGAGACAGATGGCCAGGTCCTCAGT	250
2y	241	TTAATGAAGTGGAGCCCAATGAGTTACCCCATGACTCCAGAGAACTTTACCAAGCTGTG	300
2b	251	TTAATGAAGTGGAGCCCAATGAGTTACCCCATGACTCCAGAGAACTTTACCAAGCTGTG	310
2y	301	GGTTCATGCAGAGATCCAGAGGGTAGTTATCCTGATGGATTTTCAGGCAAGCTG	360
2b	311	GGTTCATGCAGAGATCCAGAGGGTAGTTATCCTGATGGATTTTCAGGCAAGCTG	370
2y	361	CAGATAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGA	420
2b	371	CAGATAAATCCATTCATCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGA	430
2y	421	ATTATTTACTGGAGGTGCATAGCTCTTTGGTGAGAGTCTGGAGCTCCCGGGAAG	480
2b	431	ATTATTTACTGGAGGTGCATAGCTCTTTGGTGAGAGTCTGGAGCTCCCGGGAAG	490
2y	481	AATATATTCGACTCTGTGAGAAATATTAATCTCTCAGAACCCCAAGGAGTAGATTCCTAG	540
2b	491	AATATATTCGACTCTGTGAGAAATATTAATCTCTCAGAACCCCAAGGAGTAGATTCCTAG	550
Qy	541	AATGTGCAGAGAGAGTGAAGAAAGATTAATCTCTGGGTCAAGATCTCAAAACCAAGGCA	600
Db	551	AATGTGCAGAGAGAGTGAAGAAAGATTAATCTCTGGGTCAAGATCTCAAAACCAAGGCA	610
Qy	601	AAATCCCAAACTTGTACTGAGAGGTTCTGTAGATGGGATACAGAGTGGTCTGGTGA	660
Db	611	AAATCCCAAACTTGTACTGAGAGGTTCTGTAGATGGGATACAGAGTGGTCTGGTGA	670

Qy	661	ATGCTGTCTACTTCAAGGGAAGTGGAAAACTCCATTTTCAGAGAACTAAATGGGCTTT	720
Db	671	ATGCTGTCTACTTCAAGGGAAGTGGAAAACTCCATTTTCAGAGAACTAAATGGGCTTT	730
Qy	721	ATCCTTTCCGCTGTAACCTCGGCTCAGGCGACACCTGTACAGATGATGTACTTGGGTGAAA	780
Db	731	ATCCTTTCCGCTGTAACCTCGGCTCAGGCGACACCTGTACAGATGATGTACTTGGGTGAAA	790
Qy	781	AGCTAAACATTTGGATACATAGAGACCTTAAAGGCTCAGATTTCTAGAACTCCCATATGCTG	840
Db	791	AGCTAAACATTTGGATACATAGAGACCTTAAAGGCTCAGATTTCTAGAACTCCCATATGCTG	850
Qy	841	GAGATGTTAGCATGTTCTTGTTGCTTCAGATGAAATTCGGATGTGTCACATGGCTGG	900
Db	851	GAGATGTTAGCATGTTCTTGTTGCTTCAGATGAAATTCGGATGTGTCACATGGCTGG	910
Qy	901	AGCTGCTGGAAGTGAATAAATCACTATGACAACTCAACAAAGTGGACAGCAAGACAAAA	960
Db	911	AGCTGCTGGAAGTGAATAAATCACTATGACAACTCAACAAAGTGGACAGCAAGACAAAA	970
Qy	961	TGGCTGAAGATGAAGTTCGAGGTATACATACCCAGTTCAAAATTAGAAGAGCATTTATGAAC	1020
Db	971	TGGCTGAAGATGAAGTTCGAGGTATACATACCCAGTTCAAAATTAGAAGAGCATTTATGAAC	1030
Qy	1021	TCAGATCCATTTGAAAAAGCATGGGCATGGAGGACCGCTTCAACAAAGGAGCGGCCCAATT	1080
Db	1031	TCAGATCCATTTGAAAAAGCATGGGCATGGAGGACCGCTTCAACAAAGGAGCGGCCCAATT	1090
Qy	1081	TCTCAGGAGATGTCGAGAGAGGAATGACCTGTTCTTCTGAAAGTGTTCACCAAGCCCATGG	1140
Db	1091	TCTCAGGAGATGTCGAGAGAGGAATGACCTGTTCTTCTGAAAGTGTTCACCAAGCCCATGG	1150
Qy	1141	TGGATGCAATGAGAGAGCGCACTGAAGCAGCGCTGGCGACAGGAGGTGTTATGACAGGA	1200
Db	1151	TGGATGCAATGAGAGAGCGCACTGAAGCAGCGCTGGCGACAGGAGGTGTTATGACAGGA	1210
Qy	1201	GAACTGGACATGGAGGCGCCACAGTTTGTGGCAGATCATCCGTTTCTTTTCTTATATATGC	1260
Db	1211	GAACTGGACATGGAGGCGCCACAGTTTGTGGCAGATCATCCGTTTCTTTTCTTATATATGC	1270
Qy	1261	ATAAGATAACCAAGTGCATTTTATTTTCGGCAGATTTTGTCTCACCCCTAAACCTAAGGCT	1320
Db	1271	ATAAGATAACCAAGTGCATTTTATTTTCGGCAGATTTTGTCTCACCCCTAAGGCTAAGGCT	1330
Qy	1321	GCTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAGATTTGCTATTTCAAT	1380
Db	1331	GCTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAGATTTGCTATTTCAAT	1390
Qy	1381	TGCCAAAAATTTAGAGATGTTTTCTACATATTTCTCTCTCTGCTCTTCTGACCACTTCTGCTACC	1440
Db	1391	TGCCAAAAATTTAGAGATGTTTTCTACATATTTCTCTCTCTTCTGACCACTTCTGCTACC	1450
Qy	1441	CACATAATAAACAACAGAAATAATTAGACAATTGTCTATTTATTAACATGACACCCCTATT	1500
Db	1451	CACATAATAAACAACAGAAATAATTAGACAATTGTCTATTTATTAACATGACACCCCTATT	1510
Qy	1501	AATCATTTGGTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTA	1560
Db	1511	AATCATTTGGTCTTCTAAATGGGATCATGCCCATTTAGATTTTCTTACTATCAGTTTA	1570
Qy	1561	TTTTTATTAACATTAATTTTACTTTTGTATTTATTTATTTATTAATGAGTGGTGGTTTTAA	1620
Db	1571	TTTTTATTAACATTAATTTTACTTTTGTATTTATTTATTTATTAATGAGTGGTGGTTTTAA	1630
Qy	1621	ATTATTTGCTCACTGCTATTTAATGTAGCTTAATAAGTTTATAGAAGCAGATGATCTGTTA	1680
Db	1631	ATTATTTGCTCACTGCTATTTAATGTAGCTTAATAAGTTTATAGAAGCAGATGATCTGTTA	1690
Qy	1681	ATTTCCCTATTAATAAATGSCCTTTTATTTGTTCTCATAATGAAGAAATAGGTAGGTATCCCT	1740
Db	1691	ATTTCCCTATTAATAAATGSCCTTTTATTTGTTCTCATAATGAAGAAATAGGTAGGTATCCCT	1749

```
2Y 1741 CCATGCCCTTCTGTATTAATATCTGGAATAACATTAACAAATAGCAATATATGTTA 1800
Db 1750 CCATGCCCTTCTGTATTAATATCTGGAATAACATTAACAAATAGCAATATATGTTA 1809
2Y 1801 TGTGCAATTTCTAGAAATACATACACATATATATGCTGTATCTTATATTCATTTGCAAG 1860
Db 1810 TGTGCAATTTCTAGAAATACATACACATATATATGCTGTATCTTATATTCATTTGCAAG 1869
2Y 1861 TATATAATAAATAAACCTCTCTCCAAACAAC 1891
Db 1870 TATATAATAAATAAACCTCTCTCCAAACAAC 1900

RESULT 5
US-10-235-192A-45/c
; Sequence 45, Application US/10235192A
; Publication No. US20040043389A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; TITLE OF INVENTION: Methods and Compositions for Identifying
; TITLE OF INVENTION: Risk Factors for Abnormal Lipid Levels and the Diseases
; TITLE OF INVENTION: and Disorders Associated therewith
; FILE REFERENCE: WMI-011
; CURRENT APPLICATION NUMBER: US/10/235,192A
; CURRENT FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 191395
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-235-192A-45

Query Match 51.7%; Score 985.8; DB 12; Length 191395;
Best Local Similarity 99.3%; Pred. No. 1.8e-212;
Matches 990; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

2Y 902 GCTGCTGGAAGTGAATACCTATGACAACTCAACAAGTGGACCAAGCAAGCAAAAT 961
Db 165134 GCAGCTGGAAGTGAATACCTATGACAACTCAACAAGTGGACCAAGCAAGCAAAAT 165075
2Y 962 GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAAGCAATTTAGAACT 1021
Db 165074 GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAAGCAATTTAGAACT 165015
2Y 1022 CAGATCCATCTGAAAAGCATGGCGATGGAGGACGCTTCAACAAGGAGCGGGCCAAATTT 1081
Db 165014 CAGATCCATCTGAGAGCATGGCGATGGAGGACGCTTCAACAAGGAGCGGGCCAAATTT 164955
2Y 1082 CTCAGGATGTCGAGAGGAATGACCTGTTTCTTCTGAAGTGTTCACCAAGCCATGGT 1141
Db 164954 CTCAGGATGTCGAGAGGAATGACCTGTTTCTTCTGAAGTGTTCACCAAGCCATGGT 164895
2Y 1142 GGATGTGAATGAGGAGGCGCACTGAAGCAGCGCTGGCACAGGAGGTGTATGACAGGAG 1201
Db 164894 GGATGTGAATGAGGAGGCGCACTGAAGCAGCGCTGGCACAGGAGGTGTATGACAGGAG 164835
2Y 1202 AACTGGACATGGAGGCCACAGTTGTGGCAGATCATCCGTTCTTTCTTATTATGCA 1261
Db 164834 AACTGGACATGGAGGCCACAGTTGTGGCAGATCATCCCTTTCTTTCTTATTATGCA 164775
2Y 1262 TAAGATAACCAATGCAATTTATTTTTCGAGATTTTGTCTACCCCTAAATTAAGCGTG 1321
Db 164774 TAAGATAACCAATGCAATTTATTTTTCGAGATTTTGTCTACCCCTAAATTAAGCGTG 164715
2Y 1322 CTGCTTCTGAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATTTGCTATTTCAAATTT 1381
Db 164714 CTGCTTCTGAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATTTGCTATTTCAAATTT 164655
2Y 1382 GCCAAAATTTAGAGATGTTTCTTACATATTTCTGCTCTCTGACAACTTCTGCTACCC 1441
Db 164654 GCCAAAATTTAGAGATGTTTCTTACATATTTCTGCTCTCTGACAACTTCTGCTACCC 164595
```

```
QY 1442 ACTAAATAAAACACAGAAATAATTAGACAAATGCTTATTATATACATGACAAACCTATTA 1501
Db 164594 ACTAAATAAAACACAGAAATAATTAGACAAATGCTTATTATATACATGACAAACCTATTA 164535
QY 1502 ATCAATTTGCTCTTAAAAATGGGATCATGCCCAATTTAGATTTTCCCTACTATCAGTTTAT 1561
Db 164534 ATCAATTTGCTCTTAAAAATGGGATCATGCCCAATTTAGATTTTCCCTACTATCAGTTTAT 164475
QY 1562 TTTTATAACATTAACATTTTACCTTTGTTTATTTATTTATTTATATATGCTGAGTTTAAAA 1621
Db 164474 TTTTATAACATTAACATTTTACCTTTGTTTATTTATTTATTTATATATGCTGAGTTTAAAA 164415
QY 1622 TTATTGCTCAGTCCCTATTAAATGAGCTTAATAAGTTATAGAGCAGATCATCTGTTAA 1681
Db 164414 TTATTGCTCAGTCCCTATTAAATGAGCTTAATAAGTTATAGAGCAGATCATCTGTTAA 164355
QY 1682 TTTCTATCTAATAAATGCTTTTAAATTTGTTCTCATATGAAGATAAATAGGTATCCCTC 1741
Db 164354 TTTCTATCTAATAAATGCTTTTAAATTTGTTCTCATATGAAGATAAATAGGTATCCCTC 164295
QY 1742 CATGCCCTTCTGTAATAAATATCTGGAATAAATCAATTAACAAATAGGCAAAATATATGTTAT 1801
Db 164294 CATGCCCTTCTGTAATAAATATCTGGAATAAATCAATTAACAAATAGGCAAAATATATGTTAT 164235
QY 1802 GTGCATTTCTAGAAATACATACATATATATGTTCTGTATCTTATATTTCAATTCGAAGT 1861
Db 164234 GTGCATTTCTAGAAATACATACATATATATGTTCTGTATCTTATATTTCAATTCGAAGT 164175
QY 1862 ATATAATAAATAAACCTGCTTCCAAACAACAAAAAT 1898
Db 164174 ATATAATAAATAAACCTGCTTCCAAACAACAAATACAT 164138
```

RESULT 6

```
US-09-962-832-252
; Sequence 252, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-252
```

```
Query Match 51.6%; Score 984.8; DB 9; Length 1199;
Best Local Similarity 99.3%; Pred. No. 1.6e-213;
Matches 989; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 902 GCTGCTGGAAGTGAATTAACCTATGCAAACTCAACAAGTGGACCAAGCAAGCAAAAT 961
Db 177 GCAGCTGGAAGTGAATTAACCTATGCAAACTCAACAAGTGGACCAAGCAAGCAAAAT 236
QY 962 GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAAGCAATTTAGAACT 1021
Db 237 GGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTTAGAAAGCAATTTAGAACT 296
QY 1022 CAGATCCATTTCAAAAGCATGGCGATGGAGGACGCTTCAACAGGAGCGGGCCAAATTT 1081
Db 297 CAGATCCATTTCAAAAGCATGGCGATGGAGGACGCTTCAACAGGAGCGGGCCAAATTT 356
QY 1082 CTCAGGATGTCGAGAGGAATGACCTGTTTCTTCTGAAGTGTTCACCAAGCCATGGT 1141
```

b	357	CTCAGGATGTCGGAGGAAAGACCTGTTCTCTTTCTGAAGTGTCCACCAAGCCATGGT	416
y	1142	GGATGTGAATGAGGAGGGGCACTGAAGCAGCCGCTGGCAGAGGAGTGTATGACAGGGAG	1201
b	417	GGATGTGAATGAGGAGGGCACGTAAGCAGCCGCTGGCAGAGGAGTGTATGACAGGGAG	476
y	1202	AACTGGACATGAGGAGGCCACAGTTCTGGCAGACATCATCGGTTCTTTTCTTATTATGCA	1261
b	477	AACTGGACATGAGGAGGCCACAGTTCTGGCAGACATCATCGTCTTTCTTTCTTATTATGCA	536
y	1262	TAAGATAACCAAGTGCAATTTATTTTTCGGCAGATTTTGTCAACCCATAAACCTAAGCGTG	1321
b	537	TAAGATAACCAAGTGCAATTTATTTTTCGGCAGATTTTCTCAACCTATAAACCTAAGCGTG	596
y	1322	CTGCTTCTGCAAAAGATTTTTGTAGATGAGCTGTGTGCTCAGAAATTCCTATTTCAAATT	1381
b	597	CTGCTTCTGCAAAAGATTTTTGTAGATGAGCTGTGTGCTCAGAAATTTCTATTTCAAATT	656
y	1382	GCCAAAAATTTAGAGATGTTTTCTACATATTTCTGCTCTCTGAAACAATCTTCGTACCC	1441
b	657	GCCAAAAATTTAGAGATGTTTTCTACATATTTCTGCTCTCTGAAACACTCTTCGTACCC	716
y	1442	ACTAAATAAAAAACAGAAATAATTAGACAATTTGTCTATATTAACATGACAACCCATTATTA	1501
b	717	ACTAAATAAAAAACAGAAATAATTAGACAATTTGTCTATATTAACATGACAACCCATTATTA	776
y	1502	ATCATTTGGTCTTTCTAAATGGGATCATGCCCAATTTAGATTTTTCCCTTACATCAGTTTAT	1561
b	777	ATCATTTGGTCTTTCTAAATGGGATCATGCCCAATTTAGATTTTTCCCTTACATCAGTTTAT	836
y	1562	TTTTATAACATTAACTTTTACTTTTGTATTATTAATTTTATATAAATGGTGAGTTTAAAA	1621
b	837	TTTTATAACATTAACTTTTACTTTTGTATTATTAATTTTATATAAATGGTGAGTTTAAAA	896
y	1622	TTATGCTCAGCTGCCCTATTTATGATGAGCTAAATAAGTTATAGAGCAGATGATCTGTAA	1681
b	897	TTATGCTCAGCTGCCCTATTTATGATGAGCTAAATAAGTTATAGAGCAGATGATCTGTAA	956
y	1682	TTTTCTTATCTAATAAATGCCCTTTAAATGTTCTCATAAATGAAGAAATGAAGTATCCCTC	1741
b	957	TTTTCTTATCTAATAAATGCCCTTTAAATGTTCTCATAAATGAAGAAATGAAGTATCCCTC	1016
y	1742	CATGCCCTTCTGTAATAATAATCTCGMAAAAAACAATTAACAATAGGCAAAATATATGTTAT	1801
b	1017	CATGCCCTTCTGTAATAATAATCTCGMAAAAAACAATTAACAATAGGCAAAATATATGTTAT	1076
y	1802	GTGCATTTCTAGAAATACATAACACATATATATGCTGTGATCTTATATTTCAAATGCAAGT	1861
b	1077	GTGCATTTCTAGAAATACATAACACATATATATGCTGTGATCTTATATTTCAAATGCAAGT	1136
y	1862	ATATAATAATAAACCTGCTTCCAAACAAACAAAAA	1897
b	1137	ATATAATAATAAACCTGCTTCCAAACAAACATAACA	1172

RESULT 7

JS-09-813-358-178/c

Sequence 178, Application US/09813358

Patent No. US20020048759A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Pyle, Ruth

APPLICANT: Stolk, John A

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER

FILE REFERENCE: 210121.50

;; CURRENT APPLICATION NUMBER: US/09/813,358

2001-03-21

; NUMBER OF SEQ ID NOS: 222
; SOFTWARE: E224 CPO for Win

```

; SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0
; SEQ ID NO 178

```

; SEQ ID NO 1871
: LENGTH: 566

; HEIGHT: 360

; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-813-358-178

Query Match 29.4%; Score 560; DB 9; Length 566;
Best Local Similarity 100.0%; Pred. No. 4.1e-117;

Qy	1327	TCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAGAAATGTCTATTTCAAATGGCCAA	1386
Db	566	TCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAGAAATGTCTATTTCAAATGGCCAA	507
Qy	1387	AAATTTAGAGATGTTTTCTPACATATTTCTGCTCTCTCGAAACAACTTCTGTACCCACTAA	1446
Db	506	AAATTTAGAGATGTTTTCTPACATATTTCTGCTCTCTCGAAACAACTTCTGTACCCACTAA	447
Qy	1447	ATPAAAAACAGAAATAATTAGCAATTTGCTATTATTAACAATGACCAACCCCTATTAAATCAT	1506
Db	446	ATPAAAAACAGAAATAATTAGCAATTTGCTATTATTAACAATGACCAACCCCTATTAAATCAT	387
Qy	1507	TTGGTCTTCTAAAATGGGATCATGCCCATTAGATTTTTCTTACTATCAGTTTTATTTTTTA	1566
Db	386	TTGGTCTTCTAAAATGGGATCATGCCCATTAGATTTTTCTTACTATCAGTTTTATTTTTTA	327
Qy	1567	TAACAATTAACCTTTTACTTTGTTATTTTATTTTTATATAATGGTGAGTTTTTAAATTAAT	1626
Db	326	TAACAATTAACCTTTTACTTTGTTATTTATTTTTATATAATGGTGAGTTTTTAAATTAAT	267
Qy	1627	GCTCACTGCCTATTAAATGACTAATAAGTTTATAGAAGCAGATGATCTGTAAATTTCC	1686
Db	266	GCTCACTGCCTATTAAATGACTAATAAGTTTATAGAAGCAGATGATCTGTAAATTTCC	207
Qy	1687	TATCTAAATAATGCCTTTAAATGTTCTCATAAATGAAGAAATGATAGGTATCCCTCCATGC	1746
Db	206	TATCTAAATAATGCCTTTAAATGTTCTCATAAATGAAGAAATGATAGGTATCCCTCCATGC	147
Qy	1747	CCTTCTGTAATAAATATCTGGAATAAACATATAACCAATAGGCCAAATATATGTTATGTGCA	1806
Db	146	CCTTCTGTAATAAATATCTGGAATAAACATATAACCAATAGGCCAAATATATGTTATGTGCA	87
Qy	1807	TTTCTAGAAATACATAACACATATATATGCTGTATCTTATATTCAATTCGAAGTATATA	1866
Db	86	TTTCTAGAAATACATAACACATATATATGCTGTATCTTATATTCAATTCGAAGTATATA	27
Qy	1867	ATAATAAACCTGCTTCCAA	1886
Db	26	ATAATAAACCTGCTTCCAA	7

RESIST. T. a

US-09-997-279-178/c

; Sequence 178, Application US/09997279

; Publication No. US20030059781A1

; GENERAL INFORMATION:

APPLICANT: Chenault, Ruth A.

; APPLICANT: Xu, Jiangchun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER

; FILE REFERENCE: 210121.501C1

; CURRENT APPLICATION NUMBER: US/09/997,279

; CURRENT FILING DATE: 2001-11-28

; NUMBER OF SEQ ID NOS: 230

```
; SOFTWARE: FastSEQ for Windows Version 4.0
```

; SEQ ID NO 178

995 : HIGNET ;
LENGT: 566

TYPE: DNA

;
ORGANISM: Homo sapien
US-09-997-279-178

8/T-617-166-60-50

Answer Match 20 A2: score 560: DB 10: length 566:

Query Match
Best local similarity 100.0%; Pred. No. 41e-117;
23.4%; SCORE 360; LB IV; Fengch 300;

BEST LOCAL SIMILARITY 100.0%, FREQ. NO. 4:1E-117,
Matches 560: Conservative 0: Mismatches 0: Indels

MacColles 300, COMBET VACTVC U, PITSMAUCCUCUB U, ++++++U

```

QY 1327 TCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAGAAATTCGTATTTCAAAATGGCAA 1386
Db 566 TCTGCAAAAGATTTTGTAGATGAGCTGTGTGCTCAGAAATTCGTATTTCAAAATGGCAA 507
QY 1387 AAATTTAGAGATGTTTCTACATATTTCTGCTCTCTCTGAAACAACTTCTGCTACCCACTAA 1446
Db 506 AAATTTAGAGATGTTTCTACATATTTCTGCTCTCTCTGAAACAACTTCTGCTACCCACTAA 447
QY 1447 ATAAATAACACAGAAATTAATAGACAATTTGCTATTATTAATGATGACCAACCTTATTAATCAT 1506
Db 446 ATAAATAACACAGAAATTAATAGACAATTTGCTATTATTAATGATGACCAACCTTATTAATCAT 387
QY 1507 TTGGCTCTTAAATGGGATCATGCCAATTTAGATTTTCTTACTATCATCAGTTTATTTTAA 1566
Db 386 TTGGCTCTTAAATGGGATCATGCCAATTTAGATTTTCTTACTATCATCAGTTTATTTTAA 327
QY 1567 TAAACATTAACCTTTTACTTTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1626
Db 326 TAAACATTAACCTTTTACTTTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 267
QY 1627 GCTCAGTCTTATTTATGATGCTTAATAAGTTTATAGAGCAGATGATCTGTTAATTTCC 1686
Db 266 GCTCAGTCTTATTTATGATGCTTAATAAGTTTATAGAGCAGATGATCTGTTAATTTCC 207
QY 1687 TATCTAATAATGCTTTAATTTCTCATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1746
Db 206 TATCTAATAATGCTTTAATTTCTCATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 147
QY 1747 CCTTCGTAAATAATCTGGAATAAATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1806
Db 146 CCTTCGTAAATAATCTGGAATAAATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 87
QY 1807 TTCTAGAAATACATACACATATATATGCTGATCTTATTTATTTATTTATTTATTTATTTA 1866
Db 86 TTCTAGAAATACATACACATATATATGCTGATCTTATTTATTTATTTATTTATTTATTTA 27
QY 1867 ATAAATPAAACCTCTCTCCAA 1886
Db 26 ATAAATPAAACCTCTCTCCAA 7

RESULT 9
US-09-813-358-66
; Sequence 66, Application US/09813358
; Patent No. US20020048759A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Pyle, Ruth
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.501
; CURRENT APPLICATION NUMBER: US/09/813.358
; NUMBER OF SEQ ID NOS: 222
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 503
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-813-358-66

Query Match 26.1%; Score 498.8; DB 9; Length 503;
Best Local Similarity 99.6%; Pred. No. 3.2e-103;
Matches 500; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 TTGCCCGTCAGACGCAACTCAGAGAATAACAGAGAACAACAGATTGAAACAATGGAG 68
Db 2 TCAGACGTCAGACGCAACTCAGAGAATAACAGAGAACAACAGATTGAAACAATGGAG 61
QY 69 GATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCAAAAGCA 128
Db 62 GATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCAAAAGCA 121
QY 129 AGCCCAACCCAGAACCTCTTCTCTCCCATGGAGCATCTGTCACCATGGCCATGGTC 188
Db 122 AGCCCAACCCAGAACCTCTTCTCTCCCATGGAGCATCTGTCACCATGGCCATGGTC 181
QY 189 TACATGGGCTCCAGGGGCGAGCACCGAAGACCCAGATGGCCAGGTCCTTCAATTTAATGAA 248
Db 182 TACATGGGCTCCAGGGGCGAGCACCGAAGACCCAGATGGCCAGGTCCTTCAATTTAATGAA 241
QY 249 GTGGGAGCCAAATGCACTTACCCCATGACTTCCAGAGAACTTTACAGCTGTGGGTTTCATG 308
Db 242 GTGGGAGCCAAATGCACTTACCCCATGACTTCCAGAGAACTTTACAGCTGTGGGTTTCATG 301
QY 309 CAGCAGATCCAGAGGAGTATCTGATGCGATTTTTCAGGACCAAGCTGCAGATATA 368
Db 302 CAGCAGATCCAGAGGAGTATCTGATGCGATTTTTCAGGACCAAGCTGCAGATATA 361
QY 369 ATCCATTTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTTATTA 428
Db 362 ATCCATTTCATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTTATTA 421
QY 429 CTGGAAAGTCAATTAAGCTGTTTGTGAGAAGTCTCGAGCTTCCGGGAAGATATATT 488
Db 422 CTGGAAAGTCAATTAAGCTGTTTGTGAGAAGTCTCGAGCTTCCGGGAAGATATATT 481
QY 489 CGACTCTGTCAAGAAATTTACT 510
Db 482 CGACTCTGTCAAGAAATTTACT 503

RESULT 10
US-09-997-279-66
; Sequence 66, Application US/09997279
; Publication No. US20030059781A1
; GENERAL INFORMATION:
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.501C1
; CURRENT APPLICATION NUMBER: US/09/997.279
; CURRENT FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 503
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-997-279-66

Query Match 26.1%; Score 498.8; DB 10; Length 503;
Best Local Similarity 99.6%; Pred. No. 3.2e-103;
Matches 500; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 TTGCCCGTCAGACGCAACTCAGAGAATAACAGAGAACAACAGATTGAAACAATGGAG 68
Db 2 TCAGACGTCAGACGCAACTCAGAGAATAACAGAGAACAACAGATTGAAACAATGGAG 61
QY 69 GATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCAAAAGCA 128
Db 62 GATCTTTGTGTGGCAACACACTCTTTGCCCTCAATTTATTTCAAGCATCTGGCAAAAGCA 121
QY 129 AGCCCAACCCAGAACCTCTTCTCTCCCATGGAGCATCTGTCACCATGGCCATGGTC 188
Db 122 AGCCCAACCCAGAACCTCTTCTCTCCCATGGAGCATCTGTCACCATGGCCATGGTC 181
QY 189 TACATGGGCTCCAGGGGCGAGCACCGAAGACCCAGATGGCCAGGTCCTTCAATTTAATGAA 248
Db 182 TACATGGGCTCCAGGGGCGAGCACCGAAGACCCAGATGGCCAGGTCCTTCAATTTAATGAA 241
QY 249 GTGGGAGCCAAATGCACTTACCCCATGACTTCCAGAGAACTTTACAGCTGTGGGTTTCATG 308
Db 242 GTGGGAGCCAAATGCACTTACCCCATGACTTCCAGAGAACTTTACAGCTGTGGGTTTCATG 301

```

```

309 CAGCAGATCCAGAGGTAGTTATCTCGATGCGATTTTGCAGGCACAAAGCTGCAGATAA 368
302 CAGCAGATCCAGAGGTAGTTATCTCGATGCGATTTTGCAGGCACAAAGCTGCAGATAA 361
369 ATCCATTATCTCTCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGAATTATTTA 428
362 ATCCATTATCTCTCGCTCTCTCAGCTCTGCAATCAATGCATCCACAGGGAATTATTTA 421
429 CTGGAAGTGTCAATAAGTGTGTTGTTGAGAGTCTGCGAGCTTCCGGGAAGATATATT 488
422 CTGGAAGTGTCAATAAGTGTGTTGTTGAGAGTCTGCGAGCTTCCGGGAAGATATATT 481
489 CGACTCTGTCAAGAAATATTAAT 510
482 CGACTCTGTCAAGAAATATTAAT 503

```

```

RESULT 11
US-09-864-761-13940/c
; Sequence 13940, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 4917
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13940
; LENGTH: 482
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009802.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
US-09-864-761-13940
Query Match 24.9%; Score 475.6; DB 9; Length 482;
Best Local Similarity 99.2%; Pred. No. 5.9e-98;
Matches 478; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 961 TGGCTGAAGATCAAGTTGAGGTATACATACCCAGTTCAAAATTAGAAGAGATTATGAAC 1020
DB 482 TGGCTGAAGATCAAGTTGAGGTATACATACCCAGTTCAAAATTAGAAGAGATTATGAAC 423
QY 1021 TCAGATCCATTCTGAAAAGCATGGGCATGGAGACGCTTCAACAAGGAGCGGCCAATT 1080
DB 422 TCAGATCCATTCTGAGAAGCATGGGCATGGAGACGCTTCAACAAGGAGCGGCCAATT 363
QY 1081 TCTCAGGATGTCCGAGAGAGGAATGACCTGTTCTTTCTGAAGTGTTCACCAAGCCATGG 1140
DB 362 TCTCAGGATGTCCGAGAGAGGAATGACCTGTTCTTTCTGAAGTGTTCACCAAGCCATGG 303
QY 1141 TGGATGTGAATGAGGAGGCGCCTGAAGCAGCCGCTGCACAGGAGGTGTTATGACAGGA 1200
DB 302 TGGATGTGAATGAGGAGGCGCCTGAAGCAGCCGCTGCACAGGAGGTGTTATGACAGGA 243
QY 1201 GAATCGGACATGGAGGCCACAGTTTGTGCGAGATCATCCGTTTCTTTCTTATATATGC 1260
DB 242 GAATCGGACATGGAGGCCACAGTTTGTGCGAGATCATCCGTTTCTTTCTTATATATGC 183
QY 1261 ATAAGATAACCAAGTGCAATTTATTTTTCGGCAGATTTTGTCTACCCCTAAAACTAAGCGT 1320
DB 182 ATAAGATAACCAAGTGCAATTTATTTTTCGGCAGATTTTCTCTACCCCTAAAACTAAGCGT 123
QY 1321 GCTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTCAAAT 1380
DB 122 GCTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGCTCAGAAATGCTATTTCAAAT 63
QY 1381 TGCCAAAATTTAGAGATTTTCTACATATTTCTGCTCTTCTGAAACAACCTTCTGCTACC 1440
DB 62 TGCCAAAATTTAGAGATTTTCTACATATTTCTGCTCTTCTGAAACAACCTTCTGCTACC 3
QY 1441 CA 1442
DB 2 CA 1
RESULT 12
US-10-029-386-13440
; Sequence 13440, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13440
; LENGTH: 596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR18.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.32
; OTHER INFORMATION: NT HIT: M31551.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P05120, EVALUATE 1.00e-56

```



```

; OTHER INFORMATION: EST_HUMAN HIT: BF666685.1, EVALUE 0.00e+00
US-10-029-386-13440

Query Match      20.3%; Score 387.6; DB 14; Length 596;
Best Local Similarity 99.0%; Pred. No. 6.8e-78;
Matches 390; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 902 GCTGCTGGAAGTGAATAACCTATGACAACTCAACAAGTGGACCAAGACAAAAT 961
DB 203 GCAGCTGGAAGTGAATAACCTATGACAACTCAACAAGTGGACCAAGACAAAAT 262

QY 962 GCTGGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGAGCATTTATGA 1021
DB 263 GCTGGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGAGCATTTATGA 322

QY 1022 CAGATCCATCTCGAAAGCATGGGATGAGGAGCGCTTCAACAAGGACGGGCCAATTT 1081
DB 323 CAGATCCATCTCGAAAGCATGGGATGAGGAGCGCTTCAACAAGGACGGGCCAATTT 382

QY 1082 CTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATG 1141
DB 442 CTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATG 1141

QY 1142 CGATGTGAATGAGAGGCGCTGAGCAGCGCGCTGGACAGAGGTGTATGACAGGGAG 1201
DB 443 GATGTGAATGAGAGGCGCTGAGCAGCGCGCTGGACAGAGGTGTATGACAGGGAG 502

QY 1202 AACTGACATGAGAGGCCACAGTTGTGGCAGATCATCCGTTCTTTCTTTATTTATGCA 1261
DB 503 AACTGACATGAGAGGCCACAGTTGTGGCAGATCATCCCTTTCTTTCTTTATTTATGCA 562

QY 1262 TAAGATAACCAAGTGCATTTATTTTTCGGGAGA 1295
DB 563 TAAGATAACCAAGTGCATTTATTTTTCGGGAGA 596

RESULT 13
US-10-027-632-122703/c
; Sequence 122703, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122703
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-122703

Query Match      18.4%; Score 350.2; DB 15; Length 1256;
Best Local Similarity 99.2%; Pred. No. 3.4e-69;
Matches 352; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 902 GCTGCTGGAAGTGAATAACCTATGACAACTCAACAAGTGGACCAAGACAAAAT 961
DB 355 GCAGCTGGAAGTGAATAACCTATGACAACTCAACAAGTGGACCAAGACAAAAT 296

QY 962 GCTGGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGAGCATTTATGA 1021
DB 295 GCTGGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGAGCATTTATGA 236

QY 1022 CAGATCCATCTCGAAAGCATGGGATGAGGAGCGCTTCAACAAGGACGGGCCAATTT 1081
DB 235 CAGATCCATCTCGAAAGCATGGGATGAGGAGCGCTTCAACAAGGACGGGCCAATTT 176

QY 1082 CTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATG 1141
DB 175 CTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATG 116

QY 1142 GATGTGAATGAGAGGCGCTGAGCAGCGCGCTGGACAGAGGTGTATGACAGGGAG 1201
DB 115 GATGTGAATGAGAGGCGCTGAGCAGCGCGCTGGACAGAGGTGTATGACAGGGAG 56

QY 1202 AACTGACATGAGAGGCCACAGTTGTGGCAGATCATCCGTTCTTTCTTTATTT 1256
DB 55 AACTGACATGAGAGGCCACAGTTGTGGCAGATCATCCCTTTCTTTCTTTATTT 1

RESULT 14
US-10-029-386-27140
; Sequence 27140, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; POLYMORPHISM ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 27140
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR18.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.32
; OTHER INFORMATION: EST_HUMAN HIT: BF666685.1, EVALUE 0.00e+00
; OTHER INFORMATION: NT HIT: M24657.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P05120, EVALUE 3.00e-47
US-10-029-386-27140

Query Match      18.0%; Score 343.2; DB 14; Length 354;
Best Local Similarity 99.1%; Pred. No. 6.3e-68;
Matches 345; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 902 GCTGCTGGAAGTGAATAACCTATGACAACTCAACAAGTGGACCAAGACAAAAT 961
DB 7 GCAGCTGGAAGTGAATAACCTATGACAACTCAACAAGTGGACCAAGACAAAAT 66

QY 962 GCTGGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGAGCATTTATGA 1021
DB 67 GCTGGAAGATGAAGTTGAGGTATACATACCCAGTTCAAATTAGAAGAGCATTTATGA 126

QY 1022 CAGATCCATCTCGAAAGCATGGGATGAGGAGCGCTTCAACAAGGACGGGCCAATTT 1081
DB 127 CAGATCCATCTCGAAAGCATGGGATGAGGAGCGCTTCAACAAGGACGGGCCAATTT 186

QY 1082 CTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATG 1141
DB 187 CTCAGGGATGTCGGAGAGGAATGACCTGTTTCTTCTGAAAGTGTTCACCAAGCCATG 246

QY 1142 GATGTGAATGAGAGGCGCTGAGCAGCGCGCTGGACAGAGGTGTATGACAGGGAG 1201

```

247 GGATGTGAATGAGGAGGACCTGAGCAGCCGCTGCGACAGGAGGTGTATGACAGGGAG 306
1202 AACTGGACATGAGGCCACACAGTTTGTGCGAGATCATCCGTTTCTTTT 1249
307 AACTGGACATGAGGCCACACAGTTTGTGCGAGATCATCCGTTTCTTTT 354

RESULT 15
JS-10-029-386-865/c
; Sequence 865, Application US/10029386
; Publication No US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AECOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 865
; LENGTH: 544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009802.13
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.84
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.86
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.33
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: NT HIT: M1551.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF66685.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P05120, EVALUE 3.00e-45
US-10-029-386-865

Query Match 17.7%; Score 337.2; DB 14; Length 544;
Best Local Similarity 99.1%; Pred. No. 1.9e-56;
Matches 339; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 902 GCTGCTGGAAGTGAATTAACCTATGACAACTCAACAACTGAGCAGCAAGACCAAAAT 961
Db 342 GCAGCTGGAAGTGAATTAACCTATGACAACTCAACAACTGAGCAGCAAGACCAAAAT 283
Qy 962 GCCTGAAGATGAAGTTAGGTATACATACCCAGTTCAAATTAGAAGACATTATGAAT 1021
Db 282 GCCTGAAGATGAAGTTAGGTATACATACCCAGTTCAAATTAGAAGACATTATGAAT 223
Qy 1022 CAGATCCATTCTGAAAGCATGGCATGGAGGACGCTTCAACAAGGACGGGCAATTT 1081
Db 222 CAGATCCATTCTGAGNAGCATGGCATGGAGGACGCTTCAACAAGGACGGGCAATTT 163
Qy 1082 CTCAGGGATGTCGGAGAGGAATGACCTGTTCTTCTGAGTGTTCACCAAGCCATGGT 1141
Db 162 CTCAGGGATGTCGGAGAGGAATGACCTGTTCTTCTGAGTGTTCACCAAGCCATGGT 103
Qy 1142 GGATGTGAATGAGGAGGGCACTGAAGCAGCCGCTGCGACAGGAGGTGTATGACAGGGAG 1201
Db 102 GGATGTGAATGAGGAGGGCACTGAAGCAGCCGCTGCGACAGGAGGTGTATGACAGGGAG 43
Qy 1202 AACTGGACATGAGGCCACACAGTTTGTGCGAGATCATCCGTT 1243
Db 42 AACTGGACATGAGGCCACACAGTTTGTGCGAGATCATCCGTT 1

Search completed: March 12, 2004, 06:13:56
Job time : 448.712 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

DM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:18:38 ; Search time 3158.9 Seconds

(without alignments)
18036.975 Million cell updates/sec

Title: US-10-084-817-27

Perfect score: 1908
Sequence: 1 gagagattccgcgcaga.....aacaaaaataaaaaaagg 1908

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hic.*

9: gb_est1.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rtd.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gsl.*

29: gb_gsl2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1241.6	65.1	1248	29 AY416313	AY416313 Homo sapi
2	1046	54.8	1080	29 AY416314	AY416314 Pan trogl
3	1041.4	54.6	1145	9 AL550163	AL550163 AL550163
4	1027.8	53.9	1126	9 AL574114	AL574114 AL574114

5	1006.4	52.7	1965	11 AX081487	AK081487 Mus muscu
6	1000.8	52.5	1952	11 AK090049	AK090049 Mus muscu
7	970.4	50.9	1201	9 AL579236	AL579236 AL579236
8	947.6	48.9	1044	9 AL552817	AL552817 AL552817
9	932.6	48.9	991	9 AL552852	AL552852 AL552852
10	931	48.8	991	9 AL576638	AL576638 AL576638
11	925.4	48.5	1000	9 AL553610	AL553610 AL553610
12	913	47.9	1041	13 BX396200	BX396200 BX396200
13	888	46.5	1201	9 AL557133	AL557133 AL557133
14	884.6	46.4	959	9 AL544520	AL544520 AL544520
15	883.4	46.3	989	9 AL545835	AL545835 AL545835
16	880	46.1	1040	13 BX338884	BX338884 BX338884
17	873.6	45.1	897	9 AL545435	AL545435 AL545435
18	873.4	45.8	1017	9 AL547037	AL547037 AL547037
19	856.4	44.9	898	9 AL570781	AL570781 AL570781
20	854.4	44.8	1248	29 AY416315	AY416315 Mus muscu
21	840.8	44.1	1006	9 AL572149	AL572149 AL572149
22	839.8	44.0	1201	13 BX424501	BX424501 BX424501
23	834.8	43.8	901	9 AL544761	AL544761 AL544761
24	831	43.6	903	9 AL545468	AL545468 AL545468
25	805.6	42.2	833	9 AL573314	AL573314 AL573314
26	797	41.8	1201	9 AL570830	AL570830 AL570830
27	769.6	40.3	865	9 AU138758	AU138758 AU138758
28	768.6	40.3	900	13 BX424585	BX424585 BX424585
29	758.4	39.7	771	14 CD742955	CD742955 UI-H-F11
30	755.4	39.6	764	14 CD521331	CD521331 AGENCOURT
31	753.8	39.5	855	9 AU139494	AU139494 AU139494
32	753.6	39.5	764	14 CA313606	CA313606 UI-CF-FNO
33	753	39.5	818	9 AU557586	AL557586 AL557586
34	750.8	39.4	792	9 AU139617	AU139617 AU139617
35	748.2	39.2	754	13 BX099033	BX099033 BX099033
36	747.4	39.2	770	14 CA312894	CA312894 UI-CF-FNO
37	740.8	38.8	779	14 CA310353	CA310353 UI-H-F11
38	739.2	38.7	756	14 CA748486	CA748486 UI-H-F11
39	738	38.7	894	14 CD522880	CD522880 AGENCOURT
40	733.4	38.4	1201	13 BX381897	AL571400 AL571400
41	733.2	38.4	760	14 CD364488	BX381897 BX381897
42	732.8	38.4	755	14 CD743065	CD364488 UI-H-F11
43					CD743065 UI-H-F11

ALIGNMENTS

RESULT 1	AY416313	1248 bp	DNA	linear	GSS 17-DEC-2003
LOCUS	AY416313	Homo sapiens SERPINB2 gene, VIRTUAL TRANSCRIPT, partial sequence,			
DEFINITION	AY416313.1	GI:39772273			
ACCESSION	AY416313				
VERSION	AY416313.1				
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.				
	1 (bases 1 to 1248)				
	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A.,				
	Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,				
	Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.O.,				
	Adams, M.D. and Cargill, M.				
	Inferring nonneutral evolution from human-chimp-mouse orthologous				
	gene trios				
TITLE	Science 302 (5652), 1960-1963 (2003)				
JOURNAL	14671302				
PUBMED	2 (bases 1 to 1248)				
REFERENCE	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A.,				
AUTHORS	Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,				
	Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.O.,				
	Adams, M.D. and Cargill, M.				
	Direct Submission				
TITLE					

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1..1248
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1..>1248
/gene="SERPINB2"
/locus_tag="HCM5849"
ORIGIN
Query Match 65.1%; Score 1241.6; DB 29; Length 1248;
Best Local Similarity 99.7%; Pred. No. 1.7e-221;
Matches 1244; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 63 ATGGAGGATCTTTGTGGCAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGGCA 122
DB 1 ATGGAGGATCTTTGTGGCAACACACTCTTTGCCCTCAATTTATTCAAGCATCTGGCA 60
QY 123 AAAGCAAGCCCAACCCAGAACCTTCTCTCCCATGGAGCATCTGTCACCATGGCC 182
DB 51 AAAGCAAGCCCAACCCAGAACCTTCTCTCCCATGGAGCATCTGTCACCATGGCC 120
QY 183 ATGGTCTACATGGGCTCCAGGGGAGCAGCAGCAGCAGATGGCAGAGTGCTTCAGTTT 242
DB 121 ATGGTCTACATGGGCTCCAGGGGAGCAGCAGCAGCAGATGGCAGAGTGCTTCAGTTT 180
QY 243 AATGAAGTGGAGCCATGAGTATACCCCATGATGCTCCAGAGCATTTACAGCTGTGG 302
DB 181 AATGAAGTGGAGCCATGAGTATACCCCATGATGCTCCAGAGCATTTACAGCTGTGG 240
QY 303 TTATCATGAGCATCCAGAGGGTAGTTATCTCTGATGCGATTTTGCAGGCAACAGCTGCA 362
DB 241 TTATCATGAGCATCCAGAGGGTAGTTATCTCTGATGCGATTTTGCAGGCAACAGCTGCA 300
QY 363 GATATAATCATTCATCTTCCTGCTCTCAGCTCTGCAATCAATGATCCACAGGGAAT 422
DB 301 GATATAATCATTCATCTTCCTGCTCTCAGCTCTGCAATCAATGATCCACAGGGAAT 360
QY 423 TATTTACTGGAAAGTGTCAATAAGCTGTTTGGTGAGAGTCTGCGAGCTTCGGGAAGAA 482
DB 361 TATTTACTGGAAAGTGTCAATAAGCTGTTTGGTGAGAGTCTGCGAGCTTCGGGAAGAA 420
QY 483 TATTTCTGACTGTGAGAAATATTACTCTCAGAACCCAGGAGTACTTCTAGAA 542
DB 421 TATTTCTGACTGTGAGAAATATTACTCTCAGAACCCAGGAGTACTTCTAGAA 480
QY 543 TGTGAGAGAGCTAGAAAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGGCAAA 602
DB 481 TGTGAGAGAGCTAGAAAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGGCAAA 540
QY 603 ATCCCAACTTGTACTGAGGTTCTGTAGATGGGATACAGGATGGTCTCTGGTGAAT 662
DB 541 ATCCCAACTTGTACTGAGGTTCTGTAGATGGGATACAGGATGGTCTCTGGTGAAT 600
QY 663 GTGTCTACTCAAGAAAGTGAAGAACTCAATTTGAGAGAACTAAATGGGCTTTAT 722
DB 601 GTGTCTACTCAAGAAAGTGAAGAACTCAATTTGAGAGAACTAAATGGGCTTTAT 660
QY 723 CTTTCCGTGTAACTCGGCTCAGCGACACCTGTACAGATGATGTACTTGGTCAAAAG 782
DB 661 CTTTCCGTGTAACTCGGCTCAGCGACACCTGTACAGATGATGTACTTGGTCAAAAG 720
QY 783 CTAAACATTGGATACATAGAGACCTAAGGCTCAGATTTAGAGACTCCCATATGCTGGA 842
DB 721 CTAAACATTGGATACATAGAGACCTAAGGCTCAGATTTAGAGACTCCCATATGCTGGA 780
QY 843 GATGTTAGCATGTTCTTGTGCTTCCAGATGAAATTCGCGATGTGTCACCTGGCTTGGAG 902
DB 781 GATGTTAGCATGTTCTTGTGCTTCCAGATGAAATTCGCGATGTGTCACCTGGCTTGGAG 840

QY 903 CTGCTGGAAGTGAAATACCTATGACAACTCAACAGTGAGCCAGCAAGACAAAATG 962
DB 841 CTGCTGGAAGTGAAATACCTATGACAACTCAACAGTGAGCCAGCAAGACAAAATG 900
QY 963 GCTGAAGATGAAGTTGAGGTATATACATACCCAGTTCAAATTTAGAGAGCATTTGAATC 1022
DB 901 GCTGAAGATGAAGTTGAGGTATATACATACCCAGTTCAAATTTAGAGAGCATTTGAATC 960
QY 1023 AGATCCATTCTGAAAAGCATGGGCATGGAGAGCCCTTCAACAAGGGACGGGCCAATTC 1082
DB 961 AGATCCATTCTGAGAAGCATGGGCATGGAGAGCCCTTCAACAAGGGACGGGCCAATTC 1020
QY 1083 TCAGGATGTCGAGAGAGGAATGACCTGTTCTTTCTGAAGTGTTCACCAAGCCATGTTG 1142
DB 1021 TCAGGATGTCGAGAGAGGAATGACCTGTTCTTTCTGAAGTGTTCACCAAGCCATGTTG 1080
QY 1143 GATGTCAATGAGAGGGGCACTGAAGCAGCGCTGGCAGAGGTGTTATGACAGGGAGA 1202
DB 1081 GATGTCAATGAGAGGGGCACTGAAGCAGCGCTGGCAGAGGTGTTATGACAGGGAGA 1140
QY 1203 ACTGGACATGGAGGCCACAGTTTGTGGCAGATCATCGTTTCTTTCTTTATTCAT 1262
DB 1141 ACTGGACATGGAGGCCACAGTTTGTGGCAGATCATCGTTTCTTTCTTTATTCAT 1200
QY 1263 AAGATAACCAAGTGCATTTTATTTTTCGGCAGATTTTCTCACCCTAA 1310
DB 1201 AAGATAACCAAGTGCATTTTATTTTTCGGCAGATTTTCTCACCCTAA 1248
RESULT 2
AY416314 1080 bp DNA linear GSS 17-DEC-2003
LOCUS Pan troglodytes SERPINB2 gene, VIRTUAL TRANSCRIPT, partial
DEFINITION sequence, genomic survey sequence.
ACCESSION AY416314
VERSION AV416314.1 GI:39772274
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 1080)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1080)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1..1080
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
gene <1..>1080
/gene="SERPINB2"
/locus_tag="HCM5849"
ORIGIN
Query Match 54.8%; Score 1046; DB 29; Length 1080;
Best Local Similarity 97.7%; Pred. No. 5.2e-185;

Matches	1055;	Conservative	0;	Mismatches	25;	Indels	0;	Gaps	0;			
231	GTGCTT	CAGTTT	TAATGA	AGTGG	AGCCAA	TGAGTT	TACCC	CAATG	CTCCAG	AACTTT	290	
1	GTGCTG	CAGTTT	TAATGA	AGTGG	AGCCAA	TGAGTT	TACCC	CAATG	CTCCAG	AACTTT	60	
291	ACCAGCT	GTGGGTT	CATG	CAGCAG	ATCC	AGAGGG	TAGTT	TATCCT	GATGG	ATTTTGCAG	350	
61	ACCAGCT	GTGGGTT	CATG	CAGCAG	ATCC	AGAGGG	TAGTT	TATCCT	GATGG	ATTTTGCAG	120	
351	GCACAAG	CTGCAGAT	AAATCCA	ATTCAT	CTCTTCCG	CTCTCT	CACTCT	CTGCA	ATCAATGCA		410	
121	GCACAAG	CTGCAGAT	AAATCCA	ATTCAT	CTCTTCCG	CTCTCT	CACTCT	CTGCA	ATCAATGCA		180	
411	TCCACAGG	GAATAT	TTTACT	TGGAAG	TGTC	CAATAAG	CTGTTT	GGTGAGA	AGTCTGG	GAGC	470	
181	TCCACAGG	GAATAT	TTTACT	TGGAAG	TGTC	CAATAAG	CTGTTT	GGTGAGA	AGTCTGG	GAGC	240	
471	TTCCGG	GAAGAT	TATATT	CGACTCT	GTCA	GAATAT	TATCT	CTCAG	AAACCC	AGGCA	530	
241	TTCCGG	GAAGAT	TATATT	CGACTCT	GTCA	GAATAT	TATCT	CTCAG	AAACCC	AGGCA	300	
531	GACTTCT	CTAGA	TGTG	CAG	AAAG	CTAG	AAAAA	GATTAA	TTCT	CGGTCA	AGCTCAA	590
301	GACTTCT	CTAGA	TGTG	CAG	AAAG	CTAG	AAAAA	GATTAA	TTCT	CGGTCA	AGCTCAA	360
591	ACCAAGG	CAAAAT	CCCAACT	TTGTTAC	CTGA	AGGTTCT	GT	TAGAT	GGGAT	TAC	AGGATG	650
361	ACCAAGG	CAAAAT	CCCAACT	TTGTTAC	CTGA	AGGTTCT	GT	TAGAT	GGGAT	TAC	AGGATG	420
651	GTCCTG	GTGAAT	GCTGT	CTACTT	CAAG	GAAGTGG	AAAA	CTCCAT	TTG	AGAGAA	AACTA	710
421	GTCCTG	GTGAAT	GCTGT	CTACTT	CAAG	GAAGTGG	AAAA	CTCCAT	TTG	AGAGAA	AACTA	480
711	AATGGCT	TTATCCT	TTTC	CGTGA	AACTCG	CT	CAGG	CA	CACTGT	A	CAGATG	770
481	AATGGCT	TTATCCT	TTTC	CGTGA	AACTCG	CT	CAGG	CA	CACTGT	A	CAGATG	540
771	TTGCGT	GAAAAG	CTAA	CA	TTTGG	ATACAT	AGA	AGCCT	TA	AAAGGCT	CAGAT	830
541	TTGCGT	GAAAAG	CTAA	CA	TTTGG	ATACAT	AGA	AGCCT	TA	AAAGGCT	CAGAT	600
831	CCATAT	GCTGG	AGAT	TAG	CA	TGTTCT	TGCT	CCAG	ATCA	AAATG	CCGATG	890
601	CCATAT	GCTGG	AGAT	TAG	CA	TGTTCT	TGCT	CCAG	ATCA	AAATG	CCGATG	660
891	ACTGGCT	TGGAG	CTGCT	GGAAG	TGA	AAAT	TA	ACCT	ATG	CA	AAAT	950
661	ACTGGCT	TNNNN	CTGGA	AGTGA	AAAT	TA	ACCT	ATG	CA	AAAT	950	
951	AAAGCA	AAATG	CTG	CA	AGAT	GA	AGT	TTG	AGT	TTG	AGT	1010
721	AAAGCA	AAATG	CTG	CA	AGAT	GA	AGT	TTG	AGT	TTG	AGT	780
1011	CATTAT	GAACT	CAG	ATC	CAAT	CTT	G	AAAG	CAT	GG	CA	1070
781	CATTAT	GAACT	CAG	ATC	CAAT	CTT	G	AAAG	CAT	GG	CA	840
1071	CGGGCC	AAATTT	CT	CAGG	ATG	TCGG	AG	AA	TGA	CTG	TT	1130
841	CGGGCC	AAATTT	CT	CAGG	ATG	TCGG	AG	AA	TGA	CTG	TT	900
1131	CAAGCC	ATGGT	GGAT	GTG	AA	TGAGG	GG	CA	CTG	AA	AGC	1190
901	CAAGCC	ATGGT	GGAT	GTG	AA	TGAGG	GG	CA	CTG	AA	AGC	960
1191	ATGACA	GGG	AG	AACT	CG	AG	CC	CA	AGT	TTG	TG	1250
961	ATGACA	GGG	AG	AACT	CG	AG	CC	CA	AGT	TTG	TG	1020
1251	CTTATT	TG	CA	TAA	GA	TA	CA	TTT	TT	TCG	CA	1310
1021	CTTATT	TG	CA	TAA	GA	TA	CA	TTT	TT	TCG	CA	1080

RESULT 3
AL550163
LOCUS
DEFINITION
AL550163 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0D1040YG06 5-PRIME, mRNA sequence.
ACCESSION
AL550163
VERSION
AL550163.2
KEYWORDS
GI:31271981
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1. (bases 1 to 1145)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
On Feb 15, 2001 this sequence version replaced gi:12886866.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9202.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1040BD03QPI&cluster=9202.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1040BD03QPI.
Location/Qualifiers
1..1145
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1040YG06"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source

sites of the POWSFOK1 & vector. Library was normalized.									
ORIGIN	Query Match	54.6%;	Score 1041.4;	DB 9;	Length 1145;				
	Best Local Similarity	98.2%;	Pred. No. 3.7e-184;						
	Matches 1066;	Conservative	7;	Mismatches	9;	Indels	3;	Gaps	2;
Qy	10	TGCCCGT	CAGACAGCAACT	CTCAGAGAT	TAACCCAGAGAAC	CAACACAGAT	TGAACCAAT	TGGAGG	69
Db	64	TGCCCGT	CAGACAGCAACT	CAGAGAA	TAACCCAGAGAAC	CAACCCAGAT	TGAACCAAT	TGGAGG	123
Qy	70	ATCTTT	GTGTGGCAACAC	ACTCTTTTGCCCT	CAATTTATT	CAAGCATCT	GGCAAAAGCAA	129	
Db	124	ATCTTT	GTGTGGCAACAC	ACTCTTTTGCCCT	CAATTTATT	CAAGCATCT	GGCAAAAGCAA	183	
Qy	130	GGCCCA	CCCCAGAACCT	CTTTCTCT	CTCCCCAT	TGGAGCAT	CTCGTCC	CAATGGCCAT	189
Db	184	GGCCCA	CCCCAGAACCT	CTTTCTCT	CTCCCCAT	TGGAGCAT	CTCGTCC	CAATGGCCAT	243
Qy	190	ACATGG	GTCTCCAGGG	CAGCACCG	AGACCCAGAT	TGGCGAGGT	GCCTTCAGTT	TAATGAAG	249
Db	244	ACATGG	GTCTCCAGGG	CAGCACCG	AGACCCAGAT	TGGCGAGGT	GCCTTCAGTT	TAATGAAG	303
Qy	250	TGGGAG	CCCAATGAGTT	ACCCCACT	CACTCCAGAGAAC	TTTACC	CAGCTGTGGGT	TCATGC	309
Db	304	TGGGAG	CCCAATGAGTT	ACCCCACT	CACTCCAGAGAAC	TTTACC	CAGCTGTGGGT	TCATGC	363
Qy	310	AGCAGAT	CCAGAGGGT	AGTATT	ATCTGAT	TGCGAT	TTTTTG	AGGCACAAAGCT	369
Db	364	AGCAGAT	CCAGAGGGT	AGTATT	ATCTGAT	TGCGAT	TTTTTG	AGGCACAAAGCT	423


```

db 454 TCTGCTCTTCTGAACWACVTCYGCCTACCCACTAAATAAAAACACAGAAATAATTAGACAA 395
2y 1473 TTGCTCTATTATAACATGACCAACCTTATTATCATTTGGTCTTCTTAAATGGGATCATGCC 1532
db 394 TTGCTCTATTATAACATGACCAACCTTATTATCATTTGGTCTTCTTAAATGGGATCATGCC 335
2y 1533 CATTTAGATTTTCCCTACTATCATGTTATTATTTTATAACATTAACCTTTTACTTTGTTATTT 1592
db 334 CATTTAGATTTTCCCTACTATCATGTTATTATTTTATAACATTAACCTTTTACTTTGTTATTT 275
2y 1593 ATTATTTTATAAATGGTGAGTTTTTAAATATTATGCTACTGCTCTATTTAATGATAGCTAA 1652
db 274 ATTATTTTATAAATGGTGAGTTTTTAAATATTATGCTACTGCTCTATTTAATGATAGCTAA 215
2y 1653 TAAAGTTATAGAAAGCAGATGATCTGTTAATTTCCTATCTAATAAATGCCCCTTAAATGCTTC 1712
db 214 TAAAGTTATAGAAAGCAGATGATCTGTTAATTTCCTATCTAATAAATGCCCCTTAAATGCTTC 155
2y 1713 TCATAATGAAGTAATAGTAGGATCTCCCTCATGCTCTCTGTAATAAATATCTGGAATAA 1772
db 154 TCATAATGAAGTAATAGTAGGATCTCCCTCATGCTCTCTGTAATAAATATCTGGAATAA 95
2y 1773 ACATTAAACAATAGGCAATATATGTTATGTCATTTCTAGAAATATACATAACACATA 1829
db 94 ACATTAAACAATAGGCAATATATGTTATGTCATTTCTAGAAATATACATAACACATA 38

RESULT 5
LOCUS AK081487
DEFINITION AK081487 1965 bp mRNA linear HTC 20-SEP-2003
Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
library, clone:Cl30020L09 product:serine (or cysteine) proteinase
inhibitor, clade B (ovalbumin), member 2, full insert sequence.
AK081487
ACCESSION AK081487.1 GI:26349182
VERSION 1
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 2
AUTHORS Carninci, P., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitesunai, T., Tashiro, H., Itoh, M.,
Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Matsumoto, H., Sakeguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Chara, E., Watanishi, K.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5

```

```

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1965)
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
SUBMITTED (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
FEATURES
Location/Qualifiers
source
1..1965
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:Cl30020L09"
/db_xref="MGI:2413892"
/db_xref="taxon:10090"
/clone="Cl30020L09"
/tissue_type="head"
/dev_stage="16 days embryo"
/dev_stage="16 days embryo"
67..1314
/feature_id="BAC38231.1"
/protein_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/polyA_signal 1948..1953
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/polyA_site 1953
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPTGNTSMILLLLPDEIDAS
TGLELLESINFANENKWSIKDLDSDVVVYIPKFKLAQSVYELKSIQSGMEDAFN
KGLFSGMSERNDLFLSVFHQASVDVTEGVAAAGTGAVMTGRIGHGQPGQFVADH
PFLFFIMDKITDILFLVGRPSSP"
/origin 1965
/feature_id="BAC38231.1"
/translation="MEELSMTWTFALNLLKQEKSNSTONIFSPWISSTLAIVLL
GAGNTEQOMAKVLQNEIGSYGITTNRNPFSGCDFAQIQENYPSAIIQAQAGDK
IHSATSSLSSTINTPGDYLSSANKLFGESKARFKEEYIQLSKRYSTEPADVFL
CAEAREKINSWVTKTGKEIPNLLPEGSVDETKMVLVNAVYFKGKWTPEKKING
LYPFVNSHESIPVOMVPIHAKINIGVIKDLTKQILBLPT
```

Db 44 CAGAGATACGAGATTGAACAAATGGAAGACTTTCCATGGCAACACACCAATGTTGCCC 103
Qy 100 TCAATTTATTCAAGCATCTGGCAAAAGCAAGCCACCAGAACTCTCTTCTCCTCCCAT 159
Db 104 TCAATCTCTTAAGCAGATAGAAAAATCAAACTCTAGCCAGAACTCTTTATCTCTCCAT 163
Qy 160 GGAGCATCTCGTCCACCATGGCCATGCTACATGGGCTCCAGGGGAGCAGCCGAAGACC 219
Db 164 GGAGCATCTCATCAACATTTGGCCATAGTTCTCTCGGTGCTGGGGGTAACACTGNACAGC 223
Qy 220 AGATGCCCAAGGTCCTTCACTTAATGAAGTGGAGGCCAATGCAAGTACCCCACTGACTC 279
Db 224 AGATGCCCAAGGTCCTTCACTTAATGAAGTGGAGGCCAATGCAAGTACCCCACTGACTC 283
Qy 280 CAGAGAACTTTACAGCTGTGGGTTCTATGAGCAGATCCAGAAAGGTAAGTTATCTCTGATG 339
Db 284 CAGAGAACTTTACAGCTGTGGGTTCTATGAGCAGATCCAGAAAGGTAAGTTATCTCTGATG 343
Qy 340 CGATTTTGAGGCAAGCTGCGAGATAAATGCCATTCATCTCCGCTCTCTCAGCTCTG 399
Db 344 CTATTTTACAGGCACAGCAGGAGATAAATCCATTCAGCCCTCTCTCAGCTCAA 403
Qy 400 CAATCAATGCATCCACAGGGAATTTATTTACTGGAAGTGTCAATAAGCTGTTTGGTGAGA 459
Db 404 CAATCAACACACACAGGGGGAATTTTGTAGAAAGTGCAACAGCTGTTGGAGAGA 463
Qy 460 AGTCTCGAGCTTCGGGGAAGATATATTCGACTCTGTGAGAAATATTTACTCTCAGAAC 519
Db 464 AGTCTCGAAGATTCAGAAAGATATATTCGAACTCTCTAAGAAATATTTACTCAACAGAAC 523
Qy 520 CCCAGGAGTACTCTCTAGATGTCCAGAAAGCTAGAAAGGTAAGTTAAATCTCTGGG 579
Db 524 CAGAGCAGTGGACTTCTTGAATGTCTGAGAGCTAGGGAAGATTAATCTCTGGG 583
Qy 580 TCAAGACTCAACCAAGGCAAAATCCCAACTTTGTTACTCTGAGAGTTCTGTAGATGGG 639
Db 584 TCAAGACTCAACCAAGGTAAGTCCCAACTCTGTTACTCTGAGAGTTCTGTAGATGAAG 643
Qy 640 ATACCAAGGATGTCCTGTTGTAATGCTCTACTTCAAGGAAAGTGGAAACTCCATTTG 699
Db 644 ACACCAAGATGTTGCTGTGTAATGCTCTACTTCAAGGAAAGTGGAAACTCCATTTG 703
Qy 700 AGAAGAACTPAATAGGCTTTATCTCTTCCGTGTAATGCTGAGGCTCAGGCACTCTGAC 759
Db 704 AGAAGAACTPAATAGGCTTTATCTCTTCCGTGTAATGCTGAGGCTCAGGCACTCTGAC 763
Qy 760 AGATGATGACTTGGTGAAGGCTAAACATTTGGATACATAGAGACCTAAAGGCTCAGA 819
Db 764 AGATGATGTTCTCCATGCAAGGCTGAACATTTGGATACATAGAGACCTGAAGCTCAGA 823
Qy 820 TTCTAGAACTCCCATATGCTGGAGATGTTAGCATGTTCTTGTGCTTCCAGATGAATG 879
Db 824 TCCTAGAACTTCGGCATACTGGAACATACAGCATGCTCTGTTGCTTCCGATGAGATTG 883
Qy 880 CGATGTGTCCTACTGCTGAGCTGCTGGAAGTGAATAAATCACTATGACAACTCAACA 939
Db 884 AGACCACTCACTGCTGCTGGAATTTGCTGGAAGTGAATAAATCACTTCCCACTCAACA 943
Qy 940 AGTGGACCAAGCAAGCAAAATGCTGAAGATGAAGTTGAGGTATACATACCCAGTTCA 999
Db 944 AGTGGATCAGCAAGACACACTGGATGAAGATGATGTTGTGCTCTACATTTCCCAAGTTCA 1003
Qy 1000 AATTAGAAGAGCATTAATGAACCTCAGATCCATCTGAAAAGCATGGGATGAGGACGCT 1059
Db 1004 AACTGGCAACAAGCTACGAACCTCAAGTCCATTTCTCAAGCATGGGATGAGGATGCT 1063
Qy 1060 TCAACAGGACGGGCCAAATTTCTCAGGATGTCGAGAGGAATGACCTGTTCTTCTG 1119
Db 1064 TCAACAGGACGGGCCAAATTTCTCAGGATGTCGAGAGGAATGACCTTCTTCTTCTG 1123
Qy 1120 AAGTGTTCACCAAGCCATGTTGATGTAATGAGGAGGCACTGAAGCAGCGCTGGCA 1179

Db 1124 AGGTGTTCCATCAAGCCAGCGTGGATGTCACCGAGGAGGCACTGTGGCAGCTGGTGGGA 1183
Qy 1180 CAGGAGGTGTTATGACAGGAGAACTCGGACATGAGAGGCCACAGTTTGTGGCAGATCATC 1239
Db 1184 CTGGGCGAGTATGACAGGAAAGAACTGGCCATGGTGGCCACAGTTTGTGGCCGATCATC 1243
Qy 1240 CGTTTCTTTTCTTATATGACATAAGATAACCAAGTGCATTTTATTTTTCGCGAGATTT 1299
Db 1244 CTTTCTTTTCTTATCATGACAAAATTTACCCACACGATACTATTTGTGGTAGATTCT 1303
Qy 1300 GTCACCCCTAAACTAAAGCTGCTGCTTCTGCAAAAGATTTTGTAGATGAGCTGTGTC 1359
Db 1304 CTTACCCCTAAAGGGAAGACCTATTTCCACATGAGTTTGTAGCAATGAATATAAGC 1363
Qy 1360 CTCAGAAATGCTATTTCAAAATTCGCAAAATTTAGAGATGTTTCTACATATTTCTGCTC 1419
Db 1364 CTCAGAAATGCTATTTCAAGTGCCAAAATTTAAATACTTCTTACACATTTTAT --- 1419
Qy 1420 TTCTGCAACACTCTCTGCTACCCACTAAATAAAGC--ACAGAAATAATAGACAATGTC 1477
Db 1420 -----ACTTCTGCTATACACTAAATATATACCTAAAGCAATGTATACGACTTTC 1470
Qy 1478 TATTATAACATGACAAACCTTATTAATCATTTGGTCTTTCTAAATGGGATCATGCCCATTT 1537
Db 1471 AGTGCCTTACAGTATAACTCTATTAATGATTTGTTCTTAAAGTCAGATGATGTATTT 1530
Qy 1538 AGATTTTCTCTACTATCAGTTTATTTTATAACATTAACCTTTTACTTTGTTATTTATAT 1597
Db 1531 AGTTCATCTCTTATCTGCTTTGCTTTTAACTTTAGTTTTCAGTGTATTTATTT-G 1589
Qy 1598 TTTATATAATGTTGAGTTTTTAAATTTGCTCTACTGCTCTATTTAAATGAGCTAATAAG 1657
Db 1590 TTTATATAATGTTGTTTACAAATTTGTTGCTCTGTTTAAAGAACT-----GTAACA 1644
Qy 1658 TTATGAGCAGATGATCTGTTTAAATTTCTCTATCTAAATAAGTCCCTTAAATGTTCTCA 1717
Db 1645 CTACAGAGCAGAAATTAGATAATTTCTTATTAAGAAATCAGCCATTTAATTAAT 1704
Qy 1718 ATGAAGAAATAGTAGGTATCCCTCCATGCTCTCTGTAATAAATCTGGAATAAACATT 1777
Db 1705 AATGAAGGAAATAATGAGTCTTCCATCTTCCCATGATATTCACCCAGAAAAATGTA 1764
Qy 1778 AACATAGGCAATATATGTTATGTCATTTCTAGAAATCTACATAACATATATATGTC 1837
Db 1765 CTTAACAAGACATGT-----TATATCTCTATCTATATATATCATATGACGTATC 1816
Qy 1838 TGTATCTTATTTCAATTTGCAAGTATATAATAATAAACCTGCTTCCAAACA 1889
Db 1817 TGCACTACTCTATATTTAGGACTACATCATAGTAAGCATGCTTACTTACA 1868

RESULT 6

AK090049
LOCUS

DEFINITION

1952 bp mRNA linear HTC 20-SEP-2003
Mus musculus blood RCB-0035 WEHI-3 cDNA, RIKEN full-length enriched
library, clone:G430078u13 product:serine (or cysteine) proteinase
inhibitor, clade B (ovalbumin), member 2, full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

AK090049, 1 GI:26354876

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99273253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishii, K., Kikunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1952)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arawaka, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurinara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnishi, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES
source

1. .1952
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="FANTOM,DB:G430078L13"
/db_xref="MGI:2428955"
/db_xref="taxon:10090"
/clone="G430078L13"
/cell_line="RCB-0035 WEHI-3"
/tissue_type="blood"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
54. .1301
/note="unnamed protein product; putative serine (or cysteine) proteinase inhibitor, clade B

Query Match 52.5%; Score 1000.8; DB 11; Length 1952;
Best Local Similarity 73.0%; Pred. No. 1.2e-176;
Matches 1366; Conservative 0; Mismatches 477; Indels 29; Gaps 5;
QY 20 ACAGCAACTCAGAGAATAACAGAGAACCAACAGATTGAAACAATGGAGGATCTTTGTGT 79
DB 11 ATAGCAATAGTTACTATATTGAGCTGTGTAGAGGATTGAACAATGGAAGACTTTCCAT 70
QY 80 GCGAACAACACTTTTGGCCCTCAATTTTATCAAGCATCTGCGAAGAGAGAGCCCA 139
DB 71 GCGAACAACACTTTTGGCCCTCAATTTTATCAAGCATGAGAAATCAAACTACACCA 130
QY 140 GAACCTCTTCTCTCCCATCGAGCATCTCGTCCACCATGCGCATGCTTACATGGGCTC 199
DB 131 GAACATCTTTATCTCTCCATGGAGCATCTCATCAACATGGCCATAGTTCTCTCGGTGC 190
QY 200 CAGGGCAGACCGAAGACAGATGGCCAGGCTCTCAGTTTATGAAGTGGAGCCAA 259
DB 191 TGGGGTAAACACTGAACAGCAGATGGCCAAAGTCTCAGTTTATGAAGTGGAGCTTA 250
QY 260 TGCAGTTTACCCCATCAGCTCAGAGAACTTATACAGCTGTGGTTCATGCGACAGATCA 319
DB 251 TGGTATCACCACAAGAAACCCAGAGAACTTCAGTGGTGTGATTTCCGACACAGATACA 310
QY 320 GAAGGTAGTTATCTGTAGTGGATTTTGGAGGACAGACTCCAGATAAATCCATTCATC 379
DB 311 GAAGGAAATTTATCTCTAGTGTCTTTTACAGGACAGAGGAGATAAATCCATTCAGC 370
QY 380 CTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTTATTTACTGGAAAGTGT 439
DB 371 CTTCTCTCTCTCAGCTCAACATCAACACACAGGGGATTTTGTAGAAAGTGC 430
QY 440 CAATAAGCTGTTTGGTGAAGTCTGGAGCTTCCGGGAGAGATATATTGCACTGTGCA 499
DB 431 AAACAAGCTGTTTGGGAGAGAAGTCTGCAAGATTCAAGAAGAGATACATCCCAACTCTCAA 490
QY 500 GAAATATTACTCTCTCAGAACCCAGGAGTAGACTTCTAGAAATGTGCAGAGAAGTAG 559
DB 491 GAAATATTACTCAACAGACACAGAGCAGTGGACTTCTTGAATGTGTGAGAGAGTAG 550
QY 560 AAAAAAGATTATCTCTGGGTCAAGACTCAACCAAGAGGCAAAATCCCAAACTGTTTACC 619
DB 551 GGAAGAGATTATCTTGGGTCAAGACTCAACCAAGAGGTAATCCCAAACTGTTTACC 610
QY 620 TGAAGTCTCTGTAGATGGGATACAGGATGGTCTCGTGAATGTGTCTACTTCAAGG 679
DB 611 CGAAGTCTCTGTAGATGAAGACACCAAGATGGTGTGATGTGTCTACTTCAAGG 670
QY 680 AAGTGGAAATCTCCATTTGAGAAGAACTTAATGGGCTTTATCTCTTCCGTGTAACCTC 739
DB 671 AAGTGGAAATCTCCATTTGAGAAGAACTTAATGGGCTTTATCTCTTCCGTGTAACCTC 730
QY 740 GGTCTAGCAGCAGACCTGTACAGATGATGTCTTCCGTGAAAGCTTAAACATTTGGATACAT 799

polyA_signal
polyA_site
ORIGIN
Query Match 52.5%; Score 1000.8; DB 11; Length 1952;
Best Local Similarity 73.0%; Pred. No. 1.2e-176;
Matches 1366; Conservative 0; Mismatches 477; Indels 29; Gaps 5;
QY 20 ACAGCAACTCAGAGAATAACAGAGAACCAACAGATTGAAACAATGGAGGATCTTTGTGT 79
DB 11 ATAGCAATAGTTACTATATTGAGCTGTGTAGAGGATTGAACAATGGAAGACTTTCCAT 70
QY 80 GCGAACAACACTTTTGGCCCTCAATTTTATCAAGCATCTGCGAAGAGAGAGCCCA 139
DB 71 GCGAACAACACTTTTGGCCCTCAATTTTATCAAGCATGAGAAATCAAACTACACCA 130
QY 140 GAACCTCTTCTCTCCCATCGAGCATCTCGTCCACCATGCGCATGCTTACATGGGCTC 199
DB 131 GAACATCTTTATCTCTCCATGGAGCATCTCATCAACATGGCCATAGTTCTCTCGGTGC 190
QY 200 CAGGGCAGACCGAAGACAGATGGCCAGGCTCTCAGTTTATGAAGTGGAGCCAA 259
DB 191 TGGGGTAAACACTGAACAGCAGATGGCCAAAGTCTCAGTTTATGAAGTGGAGCTTA 250
QY 260 TGCAGTTTACCCCATCAGCTCAGAGAACTTATACAGCTGTGGTTCATGCGACAGATCA 319
DB 251 TGGTATCACCACAAGAAACCCAGAGAACTTCAGTGGTGTGATTTCCGACACAGATACA 310
QY 320 GAAGGTAGTTATCTGTAGTGGATTTTGGAGGACAGACTCCAGATAAATCCATTCATC 379
DB 311 GAAGGAAATTTATCTCTAGTGTCTTTTACAGGACAGAGGAGATAAATCCATTCAGC 370
QY 380 CTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTTATTTACTGGAAAGTGT 439
DB 371 CTTCTCTCTCTCAGCTCAACATCAACACACAGGGGATTTTGTAGAAAGTGC 430
QY 440 CAATAAGCTGTTTGGTGAAGTCTGGAGCTTCCGGGAGAGATATATTGCACTGTGCA 499
DB 431 AAACAAGCTGTTTGGGAGAGAAGTCTGCAAGATTCAAGAAGAGATACATCCCAACTCTCAA 490
QY 500 GAAATATTACTCTCTCAGAACCCAGGAGTAGACTTCTAGAAATGTGCAGAGAAGTAG 559
DB 491 GAAATATTACTCAACAGACACAGAGCAGTGGACTTCTTGAATGTGTGAGAGAGTAG 550
QY 560 AAAAAAGATTATCTCTGGGTCAAGACTCAACCAAGAGGCAAAATCCCAAACTGTTTACC 619
DB 551 GGAAGAGATTATCTTGGGTCAAGACTCAACCAAGAGGTAATCCCAAACTGTTTACC 610
QY 620 TGAAGTCTCTGTAGATGGGATACAGGATGGTCTCGTGAATGTGTCTACTTCAAGG 679
DB 611 CGAAGTCTCTGTAGATGAAGACACCAAGATGGTGTGATGTGTCTACTTCAAGG 670
QY 680 AAGTGGAAATCTCCATTTGAGAAGAACTTAATGGGCTTTATCTCTTCCGTGTAACCTC 739
DB 671 AAGTGGAAATCTCCATTTGAGAAGAACTTAATGGGCTTTATCTCTTCCGTGTAACCTC 730
QY 740 GGTCTAGCAGCAGACCTGTACAGATGATGTCTTCCGTGAAAGCTTAAACATTTGGATACAT 799

polyA_signal
polyA_site
ORIGIN
Query Match 52.5%; Score 1000.8; DB 11; Length 1952;
Best Local Similarity 73.0%; Pred. No. 1.2e-176;
Matches 1366; Conservative 0; Mismatches 477; Indels 29; Gaps 5;
QY 20 ACAGCAACTCAGAGAATAACAGAGAACCAACAGATTGAAACAATGGAGGATCTTTGTGT 79
DB 11 ATAGCAATAGTTACTATATTGAGCTGTGTAGAGGATTGAACAATGGAAGACTTTCCAT 70
QY 80 GCGAACAACACTTTTGGCCCTCAATTTTATCAAGCATCTGCGAAGAGAGAGCCCA 139
DB 71 GCGAACAACACTTTTGGCCCTCAATTTTATCAAGCATGAGAAATCAAACTACACCA 130
QY 140 GAACCTCTTCTCTCCCATCGAGCATCTCGTCCACCATGCGCATGCTTACATGGGCTC 199
DB 131 GAACATCTTTATCTCTCCATGGAGCATCTCATCAACATGGCCATAGTTCTCTCGGTGC 190
QY 200 CAGGGCAGACCGAAGACAGATGGCCAGGCTCTCAGTTTATGAAGTGGAGCCAA 259
DB 191 TGGGGTAAACACTGAACAGCAGATGGCCAAAGTCTCAGTTTATGAAGTGGAGCTTA 250
QY 260 TGCAGTTTACCCCATCAGCTCAGAGAACTTATACAGCTGTGGTTCATGCGACAGATCA 319
DB 251 TGGTATCACCACAAGAAACCCAGAGAACTTCAGTGGTGTGATTTCCGACACAGATACA 310
QY 320 GAAGGTAGTTATCTGTAGTGGATTTTGGAGGACAGACTCCAGATAAATCCATTCATC 379
DB 311 GAAGGAAATTTATCTCTAGTGTCTTTTACAGGACAGAGGAGATAAATCCATTCAGC 370
QY 380 CTTCCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGGAATTTATTTACTGGAAAGTGT 439
DB 371 CTTCTCTCTCTCAGCTCAACATCAACACACAGGGGATTTTGTAGAAAGTGC 430
QY 440 CAATAAGCTGTTTGGTGAAGTCTGGAGCTTCCGGGAGAGATATATTGCACTGTGCA 499
DB 431 AAACAAGCTGTTTGGGAGAGAAGTCTGCAAGATTCAAGAAGAGATACATCCCAACTCTCAA 490
QY 500 GAAATATTACTCTCTCAGAACCCAGGAGTAGACTTCTAGAAATGTGCAGAGAAGTAG 559
DB 491 GAAATATTACTCAACAGACACAGAGCAGTGGACTTCTTGAATGTGTGAGAGAGTAG 550
QY 560 AAAAAAGATTATCTCTGGGTCAAGACTCAACCAAGAGGCAAAATCCCAAACTGTTTACC 619
DB 551 GGAAGAGATTATCTTGGGTCAAGACTCAACCAAGAGGTAATCCCAAACTGTTTACC 610
QY 620 TGAAGTCTCTGTAGATGGGATACAGGATGGTCTCGTGAATGTGTCTACTTCAAGG 679
DB 611 CGAAGTCTCTGTAGATGAAGACACCAAGATGGTGTGATGTGTCTACTTCAAGG 670
QY 680 AAGTGGAAATCTCCATTTGAGAAGAACTTAATGGGCTTTATCTCTTCCGTGTAACCTC 739
DB 671 AAGTGGAAATCTCCATTTGAGAAGAACTTAATGGGCTTTATCTCTTCCGTGTAACCTC 730
QY 740 GGTCTAGCAGCAGACCTGTACAGATGATGTCTTCCGTGAAAGCTTAAACATTTGGATACAT 799

Db	731	GCATGAGACATACCTGTCCAGATGATGTTCTCTCCATGCAAGCTGAACATTTGGATACAT	790
Qy	800	AGAAGACCTAAAGGCTCAGATCTPAGAACTCCCATATGCTGGAGATGTTAGCATGTTCTT	859
Db	791	AAAGGACCTGAAGACCTCAGATCTTAGAACTCCGCATACCTGAAACATCAGCATGCTCCT	850
Qy	860	GTTCCTCCAGATGAATTCGCCATGTGTCACATGGCTTGGAGCTGCTGGAAGTGAAT	919
Db	851	GTTCCTCCGATGAGATTCAGAGCGCATCCATTCGCTTGGATTCCTGGAAGTGAAT	910
Qy	920	AACCTATGACAACTCAAAGTGGACAGCAAGACAAATAATGGCTGAAGATGAAGTTGA	979
Db	911	AAACTTTGCCAACTTCAACAAGTGGATCAGCAAGACACACTGGATGAAGATGATGTTGT	970
Qy	980	GGTATACATACCCAGTTCCAAATTTAGAAGACATTTATGAATTCAGATCCATTCGAAAG	1039
Db	971	GGTCTACATTCCTCAAGTTCAAACCTGGCACAAAGCTACGAATCTCAATTCCTCAAG	1030
Qy	1040	CATGGCGATGAGGAGCGCTTCAACAAGGAGCGGCCAAATTTCTCAGGGATGCGAGAG	1099
Db	1031	CATGGCGATGAGGATGCGCTTCAACAAGGCGAGGCCAACTTCTCAGGAATGCTGAGAG	1090
Qy	1100	GAATCACCTGTTCTTTCTGAAGTGTTCACCAAGCCATGTGTGGATGTGAATGAGGAGGG	1159
Db	1091	GAATGACCTTTTCTTTCTGAGGTGTTCCATCAAGCCAGCGTGGATGTCACCGGAGGG	1150
Qy	1160	CACGAAAGCAGCGCGCTGGGACAGGAGTGTATGACAGGGAGAACTGACATGGAGCCC	1219
Db	1151	CACGTGGCAGCTGGTGGGACTGGGCGAGTTATGACAGGAAGAACTGGCCATGGTGGCCC	1210
Qy	1220	ACAGTTTGTGGGAGATCATCCGTTCTTTTCTTTATTTATGATAAGATAACCAAGTGCAT	1279
Db	1211	ACAGTTTGTGGCGATCATCCCTTTCTTTCTTTATCATGACAAAAATTCACCACACGAT	1270
Qy	1280	TTTATTTTTCGGCAGATTTTGCTCACCCATAAACTAAGCGTGCTGCTCTGCAAAAGATT	1339
Db	1271	ACTATTGTGTGPAGATTCCTCTCACCCATAAAGGGGAAGACCTATTTCACATGAGGT	1330
Qy	1340	TTTGTAGATGAGCTGTGCGCTCAGAAATTGCTATTTTCAAATTCGCAAAAATTTAGAGATG	1399
Db	1331	TTGTAGCATGAACATATAAGCCTCAGAAATTGCATCTTCAAGTGCCAAAAGTTTAAATAC	1390
Qy	1400	TTTTCTACATATTTCTGCTCTTCTGAACAATCTCTGCTACCCACTAAATAAAAC--ACA	1457
Db	1391	TCTTACACATTTTAT-----ACTCTGCTATACCTAAATATACCTTAAAC	1437
Qy	1458	GAAATAAATAGACAAATGTCTATTATTAACATGACAAACCTATTAATCATTTGGTCTCTA	1517
Db	1438	GCAATTTGATAGCAGTCTCAGTGCTTACAGTATAACTCTATTATGTTTGTTCCTA	1497
Qy	1518	AAATGGGATATGCCCATTTAGATTTTCTTACTACTCAGTTTATTTTATACATTAAC	1577
Db	1498	AAATGATGATGATGCTATTAGTTTCATCCCTTATTACTGCTTTGCTTTATAACTTTAGTT	1557
Qy	1578	TTTACTTTGTTATTATTATTTTATAAATGGTGAGTTTAAAAATATTGCTCACCTGCCT	1637
Db	1558	TTTACAGTGTATTATT--GTTTATAAATGGTTGTTTTTACAAATGTTGCTCCTGTTT	1616
Qy	1638	ATTTAATGATCTAAATAAGTTATAGAGCAGATGATCTGTTAAATTTCTTACTATATAA	1697
Db	1617	AATGAAACT-----GTAAACACTCAGAGCGAGAAAATTAGATAATTTCTATTAAAGAA	1671
Qy	1698	TGCCTTTAATTTCTCATTAATGAAGATAAGTAGGTATCCCTCCATGCCCTTCTGTAAT	1757
Db	1672	AATCAGCCATTAAATTTAATAATGAAGGAAAAATATGAGTCTTCCATATCTTCCCATGAT	1731
Qy	1758	AAATATCTCGAAAAAACATTAACAAATGAGCAATATATGTTATGTGCAATTTCTAGAAAT	1817
Db	1732	ATTCACCCGAGAAAATGTACTTTAAACAAAGACATGT-----TATATCTCTATCAAT	1783
Qy	1818	ACATAACACATATATATGCTGTGATCTTATATTCAATTCGAAGTATATAATAATAAAC	1877
Db	1784	ATATATCATATGTACGTATCTGCAACTCATCTATAAATTAGGACTACATCATAGTAGCA	1843

Qy 1878 TGTCTCCAAACA 1889
 Db 1844 TGCTTACTTACA 1855

RESULT 7
 AL579236/c
 LOCUS AL579236 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
 DEFINITION clone CS0DH002XJ05 3-PRIME, mRNA sequence.
 ACCESSION AL579236
 VERSION AL579236.2 GI:31317520
 KEYWORDS EST.
 ORGANISM Homo sapiens (human)
 SOURCE Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 AUTHORS Li W.B., Gruber C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 16, 2001 this sequence version replaced gi:12944080.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 9202.f For more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DH002CE03NP1&cluster=9202.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DH002CE03NP1.
 Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon.9606"
 /clone="CS0DH002XJ05"
 /tissue_type="T CELLS (JURKAT CELL LINE)"
 /cell_line="JURKAT CELL LINE"
 /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo (dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN
 Query Match 50.9%; Score 970.4; DB 9; Length 1201;
 Best Local Similarity 88.7%; Pred. No. 6.3e-171;
 Matches 989; Conservative 77; Mismatches 46; Indels 3; Gaps 3;
 Qy 747 CGCACACCTGTACAGATGATGTACTTCGCGTAAAGCTTAAACATTCGATACATAGAAC 806
 Db 1113 CGSCMWCCCTTWWVRGTDITGTYTTGGGGAARAAAGAAAAAHTKGRAAAAAAARWRHC 1054
 Qy 807 CTAAGGCTCAGATT-CTAGAACTCCCATATGCTGGAGATGTTACATGCTTCTTGTTGCT 865
 Db 1053 CTAAGAGGCGCVARTTBCRGACCCCCCWTATGSKGRRWTGTWVGCAATKTYCTTGTSCT 994
 Qy 866 TCCAGATGAATTCGCCGATGTGTCCACTGGCTTCGGAGCTGCTGGAAGAGTAAATAACCTA 925
 Db 993 CCARATRAAATTGCCBATKTGCCACTGGCTTCGGAGCTGCTGGAAGAGTAAATAACCTA 934
 Qy 926 TGACAACTCAACAGTGCACGACGACAAAGACAAATGCTCAAGATGAAGTTGAGGTATA 985
 Db 933 TGACAACTCAACAAAT-GACCAAGCAAAACAAATGCTGAGATGAAGTTGAGGTAWA 875
 Qy 986 CATACCCCGAGTTCAAAATTAGAAGCAATTAATGAATCAGATCCATTCGAAAGCATGGG 1045

874 CATWCCCGAGTTCAAATTAAGAGACATTATGAATCTCAGAWCCATTCTGAGAAGCATGGG 815
 1046 CATGGAGAGCCCTTCACACAGGAGCGGCCAAATTTCTCAGGATCTCGAGAGGATGA 1105
 814 CATGGAGAGCCCTTCACACAGGAGCGGCCAAATTTCTCAGGATCTCGAGAGGATGA 755
 1106 CCTGTTTCTTTCTGAAGTGTTCACCAAGCCATGTTGGAATGTAATGAGAGGCGCACTGA 1165
 754 CCTGTTTCTTTCTGAAGTGTTCACCAAGCCATGTTGGAATGTAATGAGAGGCGCACTGA 695
 1166 AGCAGCCGCTGGCAGAGGAGTGTATGACAGGAGAGATGGAATGGAAGGCGCCACAGTT 1225
 694 AGCAGCCGCTGGCAGAGGAGTGTATGACAGGAGAGATGGAATGGAAGGCGCCACAGTT 635
 1226 TGTGGCAGATCATCGTTTCTTTCTTATTAATGATGATGATGATGATGATGATGATGAT 1285
 634 TGTGGCAGATCATCGTTTCTTTCTTATTAATGATGATGATGATGATGATGATGATGAT 575
 1286 TTTCCGCGAGATTTTCTCACCCTTAACCAATGAGCGTCTGCTCTGCAAAAGATTTTGTGA 1345
 574 TTTCCGCGAGATTTTCTCACCCTTAACCAATGAGCGTCTGCTCTGCAAAAGATTTTGTGA 515
 1346 GATGAGCTGTGCTCAGAAATGCTATTTCAATTTGCCAAAATTTAGAGATGTTTCT 1405
 514 GATGAGCTGTGCTCAGAAATGCTATTTCAATTTGCCAAAATTTAGAGATGTTTCT 455
 1406 ACATATTTCTGCTCTTCTGAACAACTTCTGTACCCACTAAATATAAAACACAGAAATAAT 1465
 454 ACATATTTCTGCTCTTCTGAACAACTTCTGTACCCACTAAATATAAAACACAGAAATAAT 395
 1466 TAGACAATTTCTATTAACATGACACACCTTAAATCAATTTGCTCTTAAATGGGA 1525
 394 TAGACAATTTCTATTAACATGACACACCTTAAATCAATTTGCTCTTAAATGGGA 335
 1526 TCATGCCCATTTAGATTTTCTTACTATCAGTTTATTTTATAACATTTAACTTTTACTTT 1585
 334 TCATGCCCATTTAGATTTTCTTACTATCAGTTTATTTTATAACATTTAACTTTTACTTT 275
 1586 GTTATTTATTTTATATAATGAGTGTGAGTTTAAATTTATGCTGCTGCTTAAATG 1645
 274 GTTATTCATTTTATATAATGAGTGTGAGTTTAAATTTATGCTGCTGCTTAAATG 215
 1646 TAGCTATAAGTTTATAGAGAGATGATCTGTAAATTTCTTAAATTTAAATGCTTTTA 1705
 214 TAGCTATAAGTTTATAGAGAGATGATCTGTAAATTTCTTAAATTTAAATGCTTTTA 155
 1706 ATTGTTCTCATAATGAAGATTAAGTAGGTATCCCTCCATGCCCTTCTGTAATAATATCT 1765
 154 ATTGTTCTCATAATGAAGATTAAGTAGGTATCCCTCCATGCCCTTCTGTAATAATATCT 95
 1766 GGAAAAACATTAACATAGGCAATATATGTTATGTCATTTCTAGAAATACATAACA 1825
 94 GGGGAARCATTTAGCAATAGGCAATATATGTTATGTCATTTCTAGAAATACATAASA 35
 1826 CATATATATGCTGCTATCTTATATTTCAATTTGCAAG 1860
 34 CATATATATGCTG-VTCITATTTTSMCMVSAAG 1

RESULT 8
 AL552817/c 990 bp mRNA linear EST 31-MAY-2003
 LOCUS
 DEFINITION
 clone CS0D1068YK02 3-PRIME, mRNA sequence.
 ACCESSION
 AL552817
 VERSION
 AL552817.2 GI:31274632
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 990)
 AUTHORS
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE
 JOURNAL
 COMMENT
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:12892068.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: secre@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9202.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0D1068BP01NPL&cluster=9202.f. Contact :
 Feng Liang Email: fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0D1068BP01NPL.
 Location/Qualifiers
 1..990
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1068YK02"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
 source

ORIGIN
 Query Match 49.7%; Score 947.6; DB 9; Length 990;
 Best Local Similarity 96.7%; Pred. No. 1.2e-166;
 Matches 957; Conservative 18; Mismatches 14; Indels 1; Gaps 1;

Qy 872 TGAATTCGCGATGTGTCACATGGCTTGGAGCTGCTGGAAGTGAATAACCTATGACAA 931
 Db 990 TGAATTCGCGATGTGTCACATGGCTTGGAGCTGCTGGAAGTGAATAACCTATGACAA 931
 Qy 932 ACTCAACAGTGCACAGCAAGAC-AAAATGCTGAAGATGAAGTTGAGGTATACATAC 990
 Db 930 ACTCAACAGTGCACAGCAAGACAAAATGCTGAAGATGAAGTTGAGGTATACATAC 871
 Qy 991 CCCAGTTCAAATTAAGAGCATTAATCACTCAGATCCATCTGAAAGCATGGCATGG 1050
 Db 870 CCCAGTTCAAATTAAGAGCATTAATCACTCAGATCCATCTGAAAGCATGGCATGG 811
 Qy 1051 AGGACGCTTCAACAGGACGCGCCAAATTTCTCAGGATGTGCGAGAGGAATGACCTGT 1110
 Db 810 AGGACGCTTCAACAGGACGCGCCAAATTTCTCAGGATGTGCGAGAGGAATGACCTGT 751
 Qy 1111 TTTCTTCTGAAGTGTCCACCAAGCCATGTCATGTGAATGAGGAGGCACTGAAGCAG 1170
 Db 750 TTTCTTCTGAAGTGTCCACCAAGCCATGTCATGTGAATGAGGAGGCACTGAAGCAG 691
 Qy 1171 CGCTGGCAGAGGAGTGTATGACAGGAGAACTGGACATGAGAGGCCACAGTTTGTGG 1230
 Db 690 CGCTGGCAGAGGAGTGTATGACAGGAGAACTGGACATGAGAGGCCACAGTTTGTGG 631
 Qy 1231 CAGATCATCCGTTCTTTTCTTATTTATGATGAAGATAACCAAGTGCATTTATTTTCG 1290
 Db 630 CAGATCATCCGTTCTTTTCTTATTTATGATGAAGATAACCAAGTGCATTTATTTTCG 571
 Qy 1291 CGAGATTTGCTCAGCCTTAAGCATGCTGCTTCTGCAAGAGATTTTGTAGATGA 1350
 Db 570 CGAGATTTTCTCAGCCTTAAGCATGCTGCTTCTGCAAGAGATTTTGTAGATGA 511
 Qy 1351 GCTGTGTGCTCAGAAATGCTATTTCAAATTTGCCAAAATTTAGAGATGTTTCTACATA 1410
 Db 510 GCTGTGTGCTCAGAAATGCTATTTCAAATTTGCCAAAATTTAGAGATGTTTCTACATA 451
 Qy 1411 TTTCTGCTTCTTGAACACTTCTGCTACCCACTTAATAAAACACAGAAATATATAGAC 1470
 Db 450 TTTCTGCTTCTTGAACACTTCTGCTACCCACTTAATAAAACACAGAAATATATAGAC 391

1471 AATTGCTATTATTAACATGACAAACCCCTATTATTAATCAATTGGTCTCTTAATAATGGATCATG 1530
 390 AATTGCTATTATTAACATGACAAACCCCTATTATTAATCAATTGGTCTCTTAATAATGGATCATG 331
 1531 CCATTAGATTTTCCCTTACTATCAGTTATTTTATTAACATTAACCTTTTACTTTGTAT 1590
 330 YCAATTTAGATTTTCCCTTACTATCAGTTATTTTATTAACATTAACCTTTTACTTTGTAT 271
 1591 TTATTATTATTATTAATGTTAGTTATTTTAAATTTATGCTCAGTCCCTATTATTAATGTTAGCT 1650
 270 TTATTATKTTATTAATGTTAGTTATTTTAAATTTTATTTTATTTTATTTTATTTTATTTAGCT 211
 1651 AATAAGTTATTAAGACAGATGATCTGTTTAAATTTTCTTATTAATTAATGCTTTTAAATGTT 1710
 210 AATAAGTTATTAAGACAGATGATCTGTTTAAATTTTCTTATTAATTAATGCTTTTAAATGTT 151
 1711 TCTCATAATGAAGATAAGTAGTATCCCTCCATGCTTCTGTTAAATAATATCTGGAAA 1770
 150 TCTCATAATGAAGATAAGTAGTATCCCTCCATGCTTCTGTTAAATAATATCTGGAAA 91
 1771 AAACATTAACATAGGCAATATATGTTATGTCATTTCTAGAAATACATAACACATAT 1830
 90 AAACATTAACATAGGCAATATATGTTATGTCATTTCTASAAATACATAACATAT 31
 1831 ATATGCTCTATCTTATTAATCAATGCAAG 1860
 30 ATATGCTCTATCTTATTAATCAATGCAAG 1

RESULT 9
 AL552852 1044 bp mRNA linear EST 31-MAY-2003
 LOCUS AL552852 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CSODI068YK02 5-PRIME, mRNA sequence.
 ACCESSION AL552852
 VERSION AL552852.2 GI:31274667
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1044)
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:12892133.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9202.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CSODI068BF01Q1&cluster=9202.f. Contact :
 Feng Liang Email: fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CSODI068BF01Q1.
 Location/Qualifiers
 1. 1044
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODI068YK02"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /notes="1st strand cDNA was primed with a NotI-oligo (dt)
 primer. Five prime end cloned, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
 source
 1. 1044
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODI068YK02"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /notes="1st strand cDNA was primed with a NotI-oligo (dt)
 primer. Five prime end cloned, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 48.9%; Score 932.6; DB 9; Length 1044;

ORIGIN

Best Local Similarity 97.9%; Pred. No. 7.3e-164;
 Matches 962; Conservative 4; Mismatches 13; Indels 4; Gaps 2;
 QY 28 TCAGAGATAAACAGAGAACCAACAGATTGAACAATGGAGGATCTTTGTGTGCAACA 87
 DB 63 TCAGAGATAAACAGAGAACCAACAGATTGAACAATGGAGGATCTTTGTGTGCAACA 122
 QY 88 CACTCTTTGGCCCTCAATTTATTCAAGCATCTGCAAAAGCAAGCCCAACCCAGAACCTCT 147
 DB 123 CACTCTTTGGCCCTCAATTTATTCAAGCATCTGCAAAAGCAAGCCCAACCCAGAACCTCT 182
 QY 148 TCCTCTCCCATGGAGCATCTGTCACCATGCCATGGTCTACATGGGCTCCAGGGGCA 207
 DB 183 TCCTCTCCCATGGAGCATCTGTCACCATGCCATGGTCTACATGGGCTCCAGGGGCA 242
 QY 208 GCACGAGACAGATGGCCAA---GGTGTTCAGTTTAAATGAATGGAGGCAATGCG 264
 DB 243 GCACGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 302
 QY 265 TTATCCCTCACTGCTCCAGAGAACTTTTACCAAGCTGTGGTTCATCCAGAGATCCAGAAG 324
 DB 303 TTATCCCTCACTGCTCCAGAGAACTTTTACCAAGCTGTGGTTCATCCAGAGATCCAGAAG 362
 QY 325 GTAGTTATCTGATGCGATTTTTCAGGCAAGCTGCAGATAAATCCATTTCCTCTCC 384
 DB 363 GTAGTTATCTGATGCGATTTTTCAGGCAAGCTGCAGATAAATCCATTTCCTCTCC 422
 QY 385 GCTCTCTAGCTCTGCAATCAATGATCCACAGGAAATTTTACTGGAAGATGTCAATA 444
 DB 423 GCTCTCTAGCTCTGCAATCAATGATCCACAGGAAATTTTACTGGAAGATGTCAATA 482
 QY 445 AGCTGTTTGGTGAGAAAGTCTGCGAGCTTCCGGGAAGAAATATATTCGACTCTGTCAAGAA 504
 DB 483 AGCTGTTTGGTGAGAAAGTCTGCGAGCTTCCGGGAAGAAATATATTCGACTCTGTCAAGAA 542
 QY 505 ATTACTCTCAGAACCCAGGAGTAGACTTCTTAGAATGTCAGAGAGAGAGAGAGAGAGAG 564
 DB 543 ATTACTCTCAGAACCCAGGAGTAGACTTCTTAGAATGTCAGAGAGAGAGAGAGAGAGAG 602
 QY 565 AGATTAAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTTTTACCTGAAG 624
 DB 603 AGATTAAATCTCTGGGTCAAGACTCAAAACCAAGGCAAAATCCCAAACTTTTACCTGAAG 662
 QY 625 GTTCTGTAGATGGGATACCAGAGATGCTGCTGGTGAATGCTGTCTACTTCAAGGAGAACT 684
 DB 663 GTTCTGTAGATGGGATACCAGAGATGCTGCTGGTGAATGCTGTCTACTTCAAGGAGAACT 722
 QY 685 GGAATCTCATTGTGAGAGAAACTAAATGGGCTTTATCTCTTCCGTGTAACTCGGCTC 744
 DB 723 GGAATCTCATTGTGAGAGAAACTAAATGGGCTTTATCTCTTCCGTGTAACTCGGCTC 782
 QY 745 AGCGCACACCTGTACAGATGATGTACTTGGTGAAAAAGCTAAACATTGGATACATAGAAG 804
 DB 783 AGCGCACACCTGTACAGATGATGTACTTGGTGAAAAAGCTAAACATTGGATACATAGAAG 842
 QY 805 ACCTAAAGGCTCAGATTCTAGACTCCCATATGCTGGAGATGTTAGCATGTTCTTTGTTGC 864
 DB 843 ACCTAAAGGCTCAGATTCTAGACTCCCATATGCTGGAGATGTTAGCATGTTCTTTGTTGC 902
 QY 865 TTCAGATGAATGGCGATGTGCTCACTGGCTGGAGCTGCTGGAAGATGAATGAATACCT 924
 DB 903 TTCAGATGAATGGCGATGTGCTCACTGGCTGGAGCTGCTGGAAGATGAATGAATACCT 962
 QY 925 ATGACAACTCAACAGTGGACCAAGCAAAAGCAAAATGGCTGAAGATGAAGTTGAGGTAT 984
 DB 963 ATGACAACTCAACAGTGGACCAAGCAAAAGCAAAATGGCTGAAGATGAAGTTGAGGTAT 1022
 QY 985 ACATACCCAGTTCAATTTAGAA 1007
 DB 1023 ACATA-CCCAGTTCAATTTAGAA 1044

RESULT 10

AL576638/c
LOCUS AL576638 991 bp mRNA linear EST 01-JUN-2003
DEFINITION Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1078YH02 3-PRIME, mRNA sequence.
ACCESSION AL576638
VERSION AL576638.2 GI:31314920
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:12938977.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9202.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgit-bin/cluster.cgi?seq=CS0D1078DD01NP1&cluster=9202.f. Contact :
Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1078DD01NP1.
Location/Qualifiers
FEATURES
source
1. 991
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1078YH02"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 48.8%; Score 931; DB 9; Length 991;
Best Local Similarity 96.7%; Pred. No. 1.5e-163;
Matches 959; Conservative 8; Mismatches 23; Indels 2; Gaps 2;
2y 884 TGTGTCCTGCTGGAGCTGCTGAAAGTGAATTAACCTATGACAACTCAACAGTG 943
Db 991 TGTGTCCTGCTGGAGCTGCTGAAAGTGAATTAACCTATGACAACTCAACAGTG 932
2y 944 GACGAGCAAGACAAATGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1003
Db 931 GACGAGCAAGACAAATGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 872
2y 1004 AGAGAGCATTTAATGACTCAGATCCATTTCTGAAAGCATGGGATGGAGGACGCTTCAA 1063
Db 871 AGAGAGCATTTAATGACTCAGATCCATTTCTGAAAGCATGGGATGGAGGACGCTTCAA 812
2y 1064 CAAGGAGCGGGCAATTTCTCAGGAGTGTGGAGAGGAATGACTGTTTCTTCTGAAGT 1123
Db 811 CAAGGAGCGGGCAATTTCTCAGGAGTGTGGAGAGGAATGACTGTTTCTTCTGAAGT 752
Qy 1124 GTTCCACCAAGCATGGTGGATGTGATGAGGAGGACCTGAGCAGCCGCTGCACAGG 1193
Db 751 -KTCACCAAGCATGGTGGATGTGATGAGGAGGACCTGAGCAGCCGCTGCACAGG 693
Qy 1194 AGGTGTTATGACAGGAGAACTGGACATGGAGGCCCAACAGTGTGTGGCAGATATCCGTT 1243
Db 692 AGGTGTTATGACAGGAGAACTGGACATGGAGGCCCAACAGTGTGTGGCAGATATCCGTT 633
Qy 1244 TCTTTTCTTATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1303
Db 632 TCTTTTCTTATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 573

Qy 1304 ACCCTAAACAAAGCGTCTGCTTCTGCAAAAGATTTTCTAGATGAGCTGTGCTCTCA 1363
Db 572 ACCCTAAACAAAGCGTCTGCTTCTGCAAAAGATTTTCTAGATGAGCTGTGCTCTCA 513
Qy 1364 GAATTGCTATTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423
Db 512 GAATTGCTATTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 453
Qy 1424 GAACAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1483
Db 452 GAACAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 393
Qy 1484 AACATGACAAACCTATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1543
Db 392 AACATGACAAACCTATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 333
Qy 1544 TCCTTACTATCAGTTTATTTTATATAACATTAACCTTTTACTTTTGTATTATTTTATA 1603
Db 332 CCCTTACTATCAGTTTATTTTATATAACATTAACCTTTTACTTTTGTATTATTTTATA 273
Qy 1604 TAATGCTGAGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTA 1663
Db 272 TAATGCTGAGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTA 213
Qy 1664 AAGCAGATGATCTGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1723
Db 212 AAGCAGATGATCTGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 153
Qy 1724 AATAAGTAGGTATCCCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1783
Db 152 AATAAGTAGGTATCCCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 93
Qy 1784 TAGGCAATATATGTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1842
Db 92 TAGGCAATATATGTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 33
Qy 1843 CTTATATCAATTCGAAGTAT 1874
Db 32 CTTATATCAATTCGAAGTAT 1

RESULT 11
LOCUS AL553610
DEFINITION AL553610 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1078YH02 5-PRIME, mRNA sequence.
ACCESSION AL553610
VERSION AL553610.2 GI:31275424
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:12893604.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9202.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgit-bin/cluster.cgi?seq=CS0D1078DD01QPI&cluster=9202.f. Contact :
Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1078DD01QPI.
Location/Qualifiers
FEATURES
source
1. 1000

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1013Y022"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 48.5%; Score 925.4; DB 9; Length 1000;
Best Local Similarity 99.8%; Pred. No. 1.6e-162;
Matches 937; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GAGGACATTCGCCGTCAGACGCACTCAGAGAAATACAGAGAAACACAGATTGAAA 60
DB 62 GATGAGCATTCGCCGTCAGACGCACTCAGAGAAATACAGAGAAACACAGATTGAAA 121
QY 61 CAATGAGGATCTTGTGTGGCAACACACTCTTGGCCCTCAATTTATTCAGCATCTGG 120
DB 122 CAATGAGGATCTTGTGTGGCAACACACTCTTGGCCCTCAATTTATTCAGCATCTGG 181
QY 121 CAAAAGCAAGCCCCACCCAGAACTCTTCCTCCCATGAGCATCTCGTCCACCATGG 180
DB 182 CAAAAGCAAGCCCCACCCAGAACTCTTCCTCCCATGAGCATCTCGTCCACCATGG 241
QY 181 CCATGCTCTACATGGGTCCTCAGGGGAGCAGCCAGAGACCATGCGCAAGGTCTCAGT 240
DB 242 CCATGCTCTACATGGGTCCTCAGGGGAGCAGCCAGAGACCATGCGCAAGGTCTCAGT 301
QY 241 TTAATGAAGTGGAGCCAAATGACATTAACCCCATGACTCCAGAGAACTTTACCATCTGG 300
DB 302 TTAATGAAGTGGAGCCAAATGACATTAACCCCATGACTCCAGAGAACTTTACCATCTGG 361
QY 301 GGTTCATGACGAGATCCAGAAAGGTAGTTATCTCGATGCGATTTTGAGGCACAAAGCTG 360
DB 362 GGTTCATGACGAGATCCAGAAAGGTAGTTATCTCGATGCGATTTTGAGGCACAAAGCTG 421
QY 361 CAGATAAATCCATTCATCTTCCTCGCTCTCCAGCTCTGCAATCAATCAATCCACAGGGA 420
DB 422 CAGATAAATCCATTCATCTTCCTCGCTCTCCAGCTCTGCAATCAATCAATCCACAGGGA 481
QY 421 ATTATTTACTGGAAGTGTCAATGAAGTGTGTTGTGAGAACTGTCGAGCTTCGGGGAAG 480
DB 482 ATTATTTACTGGAAGTGTCAATGAAGTGTGTTGTGAGAACTGTCGAGCTTCGGGGAAG 541
QY 481 AATATATTGCACTCTGTCAGAAATATTACTCTCTCAGAACCCCGAGCAGTAGACTTCCTAG 540
DB 542 AATATATTGCACTCTGTCAGAAATATTACTCTCTCAGAACCCCGAGCAGTAGACTTCCTAG 601
QY 541 AATGTGAGAAAGAGCTAGAAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGGCA 600
DB 602 AATGTGAGAAAGAGCTAGAAAAGATTAATTCCTGGGTCAAGACTCAAAACCAAGGCA 661
QY 601 AATATCCAAACTTGTACCTCAAGGTTCTGTAGATGGGGATACCAAGGATGTCCTGGTGA 660
DB 662 AATATCCAAACTTGTACCTCAAGGTTCTGTAGATGGGGATACCAAGGATGTCCTGGTGA 721
QY 661 ATGCTGTCTACTTCAAGGAAGTGGAAAATCCCAATTTGAGAGAACTTAATATGGCTTT 720
DB 722 ATGCTGTCTACTTCAAGGAAGTGGAAAATCCCAATTTGAGAGAACTTAATATGGCTTT 781
QY 721 ATCCTTTCCGTGTAAACTCGCTCAGCCGACACCTGTGATGATGATGATGATGATGATG 780
DB 782 ATCCTTTCCGTGTAAACTCGCTCAGCCGACACCTGTGATGATGATGATGATGATGATG 841
QY 781 AGCTTAACATTTGGATACATAGAGACCTTAAGGCTCAGATTCTAGAACTCCCATATGCTG 840
DB 842 AGCTTAACATTTGGATACATAGAGACCTTAAGGCTCAGATTCTAGAACTCCCATATGCTG 901
QY 841 GAGATGTAGCAT-GTTCTCTGTTGCTTCCAGATGAATTTGCCGATGTGTCTCCTGCTTG 899

```

```

DB 902 GAGATGTAGCATGTCTCTTGTGTTGCTTCCAGATGAATTCGCGATGTGCTCACTGGCTTG 961
QY 900 GAGCTGCTGGAAGTGAATTAACCTATGACAAACTCAAC 938
DB 962 GAGCTGCTGGAAGTGAATTAACCTATGACAAACTCAAC 1000

RESULT 12
BX396200
LOCUS
DEFINITION
  BX396200 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
  clone CS0D1013Y022 5-PRIME, mRNA sequence.
ACCESSION
  BX396200
VERSION
  BX396200.1 GI:30628769
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 1041)
AUTHORS
  Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE
  Full-length cDNA libraries and normalization
JOURNAL
  Unpublished (2001)
COMMENT
  Contact: Genoscope
  Genoscope - Centre National de Sequencage
  BP 191 91006 EVRY cedex - France
  Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
  Library was constructed by life technologies, a division of
  Invitrogen. This sequence belongs to sequence cluster 9202.f For
  more information about this cluster, see
  http://www.genoscope.cns.fr/
  cgi-bin/cluster.cgi?seq=CS0D1013BH1Q1&cluster=9202.f. Contact :
  Feng Liang Email: fliang@lifetech.com URL :
  http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
  Faraday Avenue Genoscope sequence ID : CS0D1013BH1Q1.
FEATURES
  Location/Qualifiers
    1..1041
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0D1013Y022"
      /tissue_type="PLACENTA COT 25-NORMALIZED"
      /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
      /notes="1st strand cDNA was primed with a NotI-oligo(dT)
      primer. Five prime end enriched, double-strand cDNA was
      digested with Not I and cloned into the Not I and EcoR V
      sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 47.9%; Score 913; DB 13; Length 1041;
Best Local Similarity 91.1%; Pred. No. 3.3e-160;
Matches 894; Conservative 68; Mismatches 18; Indels 1; Gaps 1;

QY 5 AGCATTCCTCCGTCAGACAGCACTCAGAGAAATACAGAGAAACACCAAGATTGAAACAAT 64
DB 61 AGCATTCCTCCGTCAGACAGCACTCAGAGAAATACAGAGAAACACCAAGATTGAAACAAT 120
QY 65 GGAGGATCTTGTGTGGCAACACACTCTTGGCCCTCAATTTATTCAGCATCTGCGAAA 124
DB 121 GGAGGATCTTGTGTGGCAACACACTCTTGGCCCTCAATTTATTCAGCATCTGCGAAA 180
QY 125 AGCAAGCCCAACCCAGAACCTTCTCTCTCCCATGGAGCATCTCGTCCACCATGGCCAT 184
DB 181 AGCAAGCCCAACCCAGAACCTTCTCTCTCCCATGGAGCATCTCGTCCACCATGGCCAT 240
QY 185 GGTCTCATGGGCTCCAGGGGCGAGCAGCAGAGACAGATGGCCAGGCTTCAGTTTAA 244
DB 241 GGTCTCATGGGCTCCAGGGGCGAGCAGCAGAGACAGATGGCCAGGCTTCAGTTTAA 300
QY 245 TGAAGTGGGAGCCCAATGCAGTTACCCCATGACTCCAGAGAACTTTTACCAGCTGTGGTT 304
DB 301 TGAAGTGGGAGCCCAATGCAGTTACCCCATGACTCCAGAGAACTTTTACCAGCTGTGGTT 360

```

```
305 CATGCACAGATCCAGAGGGTAGTTATCCGTGATGGGATTTTCAGGACACAGCTGAGA 364
361 CATGCACAGATCCAGAGGGTAGTTATCCGTGATGGGATTTTCAGGACACAGCTGAGA 420
365 TAAATCAATTCATCTCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAATTA 424
421 TAAATCAATTCATCTCTCCGCTCTCTCAGCTCTGCAATCAATGCAATCCACAGGGAATTA 480
425 TTTACTGGAAGTGTCAATAAGTCTTTGGTGAGAGTCTGCGAGCTTCGGGGAAGAATA 484
481 TTTACTGGAAGTGTCAATAAGTCTTTGGTGAGAGTCTGCGAGCTTCGGGGAAGAATA 540
485 TATTGCACTCTGTGAGAAATATTATCTCTCAGAACCCAGGAGTAGACTTCTAGAAATG 544
541 TATWCGACTCTGTCAAAAAAAMWATCTCABAACCCCGGAGTAGACTTCTAGAAATG 600
545 TGCAGAGAAGCTAGAAAAAGATTAATCTCTGGGTCAAGACTCAAAACAAAGGCCAAAT 604
601 TGCAGAAARMTAGAAAAAAMWKCCTGGGTCAARMTCAAAACAAAGGCCAAAT 660
605 CCCAACTTGTACTGTAAGTCTCTGTAGATGGGATACAGGATGGTCTGTGTGATGC 664
661 CCCAACTTGTWCTGWARGTCTGTAGATGGGATACAGGATGGTCTGTGTGATGC 720
665 TGTCTACTTCAAAGGAAGTGAAGAACTCAATTTGAGAGAAACTAAATGGGCTTTATCC 724
721 TGTCTACTTCAAAGGAAGTGAAGAACTCAATTTGAGAGAAACTAAATGGGCTTTATCC 780
725 TTTCCGCTGAATCGCTCAGCGCACACCTGTACAGATGATGTACTTGGTGAAGCT 784
781 TTTCCGCTGAATCGCTCAGCGCACACCTGTACAGATGATGTACTTGGTGAAGAAW 840
785 AAACATTGGATACATAGAGACCTAAAGCTCAGATTTCTAGAACTCCCATATCTGAGAA 844
841 AAAAAMWGAATAAAAMWAAAMWAAAGCTCAGATTTCTAGAACTCCCATATCTGAGAA 900
845 TGTAGCATGTTCTTGTGCTTCCAGATGAAATTCGAGTGTCCACTGGCTTGGAGCT 904
901 TGTWGAATGTTCTTGTGCTTCCAGATGAAAMWRCGATGTCCACTGGCTTGGAGCT 960
905 GCTGGAAGTGAATTAACCTATGACAACTC-AACAAGTGGACCAAGACAAATGG 963
961 GCTGGGAARWGAAMWAAAMWATGAAAMWCAACAAAGTGGGCGAGCAAGACAAATRG 1020
964 CTGAAGATGAATTTGAGGTAT 984
1021 CTGAGTGAATTTGGGTTT 1041
```

```
RESULT 13
AL557133
LOCUS
DEFINITION
AL557133 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
clone CSODH002YJ05 5-PRIME, mRNA sequence.
ACCESSION
AL557133
VERSION
AL557133.2 GI:31278933
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12900443.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
```

```
9202.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODH002CE03Q01&cluster=9202.f. Contact :
Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODH002CE03Q01.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODH002YJ05"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/cell_line="JURKAT CELL LINE"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
```

ORIGIN

```
Query Match 46.5%; Score 888; DB 9; Length 1201;
Best Local Similarity 98.1%; Pred. NO. 1.5e-155;
Matches 926; Conservative 4; Mismatches 9; Indels 5; Gaps 3;

Qy 3 GGAGCATTTGCCGTCAGACAGCAACTCAGAGAATAAACAGAGAAACACAGATTGAACA 62
Db 48 GGAGCATTTGCCGTCAGACAGCAACTCAGAGAATAAACAGAGAAACACAGATTGAACA 107

Qy 63 ATGGAGGATCTTTGTGTGGCAACACACTTTTGGCCCTCAATTTATTCAGCATCTGGA 122
Db 108 ATGGAGGATCTTTGTGTGGCAACACACTTTTGGCCCTCAATTTATTCAGCATCTGGA 167

Qy 123 AAAGCAAGCCCCACCCAGCAACCTTCTCTCCATGGAGCATCTCGTCCACCATGCC 182
Db 168 AAAGCAAGCCCCACCCAGCAACCTTCTCTCCATGGAGCATCTCGTCCACCATGCC 227

Qy 183 ATGTCTACATGGGTCCTCAGGGGCGAGCAGCAGAGACAGATGCGCAAGGTCTTCAGTT 242
Db 228 ATGTCTACATGGGTCCTCAGGGGCGAGCAGCAGAGACAGATGCGCAAGGTCTTCAGTT 287

Qy 243 AATGAAGTGGAGGCAATGCACTTACCCCATGACTCCAGAGAACTTTACAGCTGTGG 302
Db 288 AATGAAGTGGAGGCAATGCACTTACCCCATGACTCCAGAGAACTTTACAGCTGTGG 347

Qy 303 TTCTGACAGCATCCAGAAAGGTAGTTATCTCTGATGCGATTTTGCAGGACACAGCTGCA 362
Db 348 TTCTGACAGCATCCAGAAAGGTAGTTATCTCTGATGCGATTTTGCAGGACACAGCTGCA 407

Qy 363 GATAAATCCATTCATCTCTTCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGAAT 422
Db 408 GATAAATCCATTCATCTCTTCGCTCTCTCAGCTCTGCAATCAATGATCCACAGGAAT 467

Qy 423 TATTTACTGGAAGTGTCAATAAGCTTTTGGTGAGAAGTCTCGAGCTTCCGGGGAAGAA 482
Db 468 TATTTACTGGAAGTGTCAATAAGCTTTTGGTGAGAAGTCTCGAGCTTCCGGGGAAGAA 527

Qy 483 TATATTCAGCTCTGTCAGAAATATTACTCTCAGAACCCAGGAGTAGACTTCTAGAA 542
Db 528 TATATTCAGCTCTGTCAGAAATATTACTCTCAGAACCCAGGAGTAGACTTCTAGAA 587

Qy 543 TGTGCAAGAGAGCTAGAAAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAA 602
Db 588 TGTGCAAGAGAGCTAGAAAAAGATTAATCTCTGGGTCAAGACTCAAAACCAAGGCAAA 647

Qy 603 ATCCCAAACTTTTACCTGAAGTCTGTAGATGGGATACAGGATGGTCTCGTGAAT 662
Db 648 ATCCCAAACTTTTACCTGAAGTCTGTAGATGGGATACAGGATGGTCTCGTGAAT 707

Qy 663 GCTGCTACTTCAAAGGAAGTGGAAAACTCCATTTTGAAGAAACATAAATGGGCTTTAT 722
Db 708 GCTGCTACTTCAAAGGAAGTGGAAAACTCCATTTTGAAGAAACATAAATGGGCTTTAT 767
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

DN nucleic - nucleic search, using sw model
Run on: March 10, 2004, 08:38:17 ; Search time 9432.34 Seconds
(without alignments)
17011.249 Million cell updates/sec

Title: US-10-084-817-30
Perfect score: 3702
Sequence: 1 gcctagggaatgacaggcat.....aagggcggtaccgagaac 3702

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.red.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3116.4	84.2	3610	6	AX330212	AX330212 Sequence
2	3116.4	84.2	3610	6	AX336432	AX336432 Sequence
3	3116.4	84.2	3610	6	AX409760	AX409760 Sequence
4	3116.4	84.2	3610	9	HUMCDSM	M83216 Human aorta
5	3116.4	84.2	3610	11	G28707	G28707 SMS3978 Er
6	2222.6	60.0	4091	9	HSM808959	BX648808 Homo sapi
7	1570.2	42.4	2215	6	AR001263	AR001263 Sequence
8	1570.2	42.4	2215	6	E05382	E05382 DNA encodin
9	1570.2	42.4	2215	6	I23421	I23421 Sequence 14
10	1562	42.2	4130	9	BC040354	BC040354 Homo sapi
11	1558	42.1	2198	9	HUMLCAL	D90452 Homo sapien
12	1522	41.1	2375	9	HUMCALD	M64110 Human calde
13	1481.6	40.0	2137	6	AR001262	AR001262 Sequence
14	1481.6	40.0	2137	6	I23420	I23420 Sequence 13
15	1469.4	39.7	2120	9	HUMCA2	D90453 Homo sapien
16	1455.8	39.3	3342	6	AX202102	AX202102 Sequence
17	1433.6	38.7	2137	6	E05381	E05381 DNA encodin
18	1312.2	35.4	3555	4	RABRSMC	L37206 Oryctolagus
19	1116.2	30.2	4065	4	RABLCA	L37147 Oryctolagus
20	1096.8	29.6	171530	9	AC083870	AC083870 Homo sapi
21	1087.8	29.4	1677	6	AR001261	AR001261 Sequence
22	1087.8	29.4	1677	6	I23419	I23419 Sequence 12
23	1085.6	29.3	141930	9	AC146389	AC146389 Pan trogl
24	999.2	27.0	1599	6	AR001260	AR001260 Sequence
25	999.2	27.0	1599	6	I23418	I23418 Sequence 11
26	993.2	26.8	164345	9	AC145994	AC145994 Pan trogl
27	993	26.8	1896	4	AY154474	AY154474 Bos tauru
28	937.4	25.3	939	6	AR001259	AR001259 Sequence
29	937.4	25.3	939	6	I23417	I23417 Sequence 10
30	925.2	25.0	4108	5	CHKSMCA	J04968 Chicken smo
31	917.4	24.8	5541	10	RNU18419	U18419 Rattus norv
32	909	24.6	2564	5	CHKICAL	M28417 Chicken h-c
33	906.4	24.5	2831	10	BC019435	BC019435 Mus muscu
34	889	24.0	3941	9	AF247820	AF247820 Homo sapi
35	839.4	22.7	2637	9	HSM806613	BX538339 Homo sapi
36	691	18.7	705	6	AX885040	AX885040 Sequence
37	691	18.7	705	6	BD024650	BD024650 Sequence
38	689.2	18.6	756	6	BD079603	BD079603 Cancer-as
39	653.4	17.6	2696	6	HSM809239	BX649088 Homo sapi
40	597.6	16.1	706	6	BD079604	BD079604 Cancer-as
41	551.8	14.9	1126	9	BC014035	BC014035 Homo sapi
42	544.8	14.7	548	6	AX885706	AX885706 Sequence
43	544.8	14.7	548	6	BD025316	BD025316 Sequence
44	528.8	14.3	169199	9	AC090497	AC090497 Homo sapi
45	502	13.6	196487	2	AC109173	AC109173 Mus muscu

ALIGNMENTS

RESULT 1
AX330212
LOCUS AX330212 3610 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 721 from Patent WO0194629.
ACCESSION AX330212
VERSION AX330212.1 GI:18103190
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endreass, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature

QY 757 GGAAAGAGCCCAAAACGCTTCTGAGAGGCTCTGGAGCGGAGAGGAGTTGACCCAAAC 816
Db 541 GGAAAGAGCCCAAAACGCTTCTGAGAGGCTCTGGAGCGGAGAGGAGTTGACCCAAAC 600
QY 817 AATAACAGATGCAAGTCTGCTGCTCCCAAGCAGAGAAATGCAAAATGACACAGAGAGAAA 876
Db 601 AATAACAGATGCAAGTCTGCTGCTCCCAAGCAGAGAAATGCAAAATGACACAGAGAGAAA 660
QY 877 TGAACCTACCGAG 936
Db 661 TGAACCTACCGAG 720
QY 937 AACAG 996
Db 721 AACAG 780
QY 997 GAAAG 1056
Db 781 GAAAG 840
QY 1057 AGAAATCAGGTAGAGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1116
Db 841 AGAAATCAGGTAGAGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 1117 GGTAAATGCTAAATAAATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1176
Db 901 GGTAAATGCTAAATAAATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 1177 GGAACAAGGTTTCAGATCAGATTTCCCATCATGAAAGAGAGAGAGAGAGAGAGAGAGAG 1236
Db 961 GGAACAAGGTTTCAGATCAGATTTCCCATCATGAAAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1237 AGCTGAGGAG 1296
Db 1021 AGCTGAGGAG 1080
QY 1297 AGAACAAGAGATAGCAGATGAAACGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1356
Db 1081 AGAACAAGAGATAGCAGATGAAACGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1357 AGAAAG 1416
Db 1141 AGAAAG 1200
QY 1417 AGCAGAGGAG 1476
Db 1201 AGCAGAGGAG 1260
QY 1477 GATAAAGGAG 1536
Db 1261 GATAAAGGAG 1320
QY 1537 AAGGAG 1596
Db 1321 AAGGAG 1380
QY 1597 AGAGCAG 1656
Db 1381 AGAGCAG 1440
QY 1657 AAGAGGAG 1716
Db 1441 AAGAGGAG 1500
QY 1717 AGAAGATAAACCTTCAGACAGCTGTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1776
Db 1501 AGAAGATAAACCTTCAGACAGCTGTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1777 GCAAGCTTAAAG 1836
Db 1561 GCAAGCTTAAAG 1620
QY 1837 AGATGAAAGAGATTAAAG 1896

Db 1621 AGATGAAAGAGATTAAAG 1680
QY 1897 AAGAGAGAGAGATTACAGAGATTAGTCGAGAAATGGAGAAATTCATGACCCAAACTTAA 1956
Db 1681 AAGAGAGAGAGATTACAGAGATTAGTCGAGAAATGGAGAAATTCATGACCCAAACTTAA 1740
QY 1957 ACATACAGAGAAATCTTTCAGCCGCTTGGAGGAGGCGCAGCGTGGACACCAAGAGGC 2016
Db 1741 ACATACAGAGAAATCTTTCAGCCGCTTGGAGGAGGCGCAGCGTGGACACCAAGAGGC 1800
QY 2017 TGAGGCGCCGCCAGGTGGAAGCCGCAAAAGCTGGAGGAGTTTCGTCGTCGCGCG 2076
Db 1801 TGAGGCGCCGCCAGGTGGAAGCCGCAAAAGCTGGAGGAGTTTCGTCGTCGCGCG 1860
QY 2077 GGAGACCGAG 2136
Db 1861 GGAGACCGAG 1920
QY 2137 GCTGGAGGAACTCAAG 2196
Db 1921 GCTGGAGGAACTCAAG 1980
QY 2197 GAG 2256
Db 1981 GAG 2040
QY 2257 AAGAGAGAGAGATTGAAAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2316
Db 2041 AAGAGAGAGATTGAAAGGCGAG 2100
QY 2317 AGATGGCTTTCAGATCAG 2376
Db 2101 AGATGGCTTTCAGATCAG 2160
QY 2377 CAAGATAG 2436
Db 2161 CAAGATAG 2220
QY 2437 ATGACCCATCAGCAG 2496
Db 2221 ATGACCCATCAGCAG 2280
QY 2497 TGCAATTTGAGGAG 2556
Db 2281 TGCAATTTGAGGAG 2340
QY 2557 TCCTGCTGAG 2616
Db 2341 TCCTGCTGAG 2400
QY 2617 CCCACCTGACAG 2676
Db 2401 CCCACCTGACAG 2460
QY 2677 CCGCATCAATGAAATGGCTTAACTAAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2736
Db 2461 CCGCATCAATGAAATGGCTTAACTAAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
QY 2737 TTCTGACTTGGAG 2796
Db 2521 TTCTGACTTGGAG 2580
QY 2797 GGATAGAGTCACTTCCGCCACTAAAGGTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2856
Db 2581 GGATAGAGTCACTTCCGCCACTAAAGGTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640
QY 2857 ACGCAG 2916
Db 2641 ACGCAG 2700
QY 2917 ACTAAATTTGGGTTCACTTATCTTTTATTTTCAATATCCAGTAAACCCATGTTATTTATC 2976

source	1. .3610 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /tissue_type="aorta" /dev_stage="adult" join(230. .850,1616. .2611) gene CDS 230. .2611 /gene="caldesmon" /note="smooth muscle form" /codon_start=1 /product="caldesmon" /protein_id="AAA58419.1" /db_xref="GI:180195" /translation="MDDFERRELRRLRKRKEMRLAEIRAYORNDDEEEAARERRRR ARERLRKQEEESLGOVTDQVEVNAQNSVPDEEAKTTTNTQVGDDEAFLERLAR RERERLRQLERLROKQEPDPTITDASLSLPSRMNDTAENETTEKEKSSRORY EIESETVTKSYQXNDWRDASENKEDKEESEEKPKRGSIGENQVWVEEKITE SQEETVVLKNGQIISSEPKQEEEDKEEESHEEKREERDEKRAEAERLARLEAE RERIKAQDDKQIADIERAIEAEKAAQERREERAEERREEMREKGAABERQIKBE EKAAABERQIKSEKKAABERQRIKEEKKAABERQARAEEKAKVEEQRNKQL EKKRAAQETIKGEKVKQKLEKGVNKKQAQEDKLQTLAVLKKQBEKGTQVQAKREK LOEDKPTFKKEIKDEKIKKDEKPEKVEKVMDRKKFTVEKSONGFETHKLKUTEN TFSPRGGASVDVTEAEGAPOVEAKRLEELERLRREKGETSESEFEKLKQKQAALELE ELKKRERRRKVLSEERQRQBEADRLKREEEKRLKEEIERRRAEAERKQWPE DASDDKKPKFTPKGSLKLEAEFLNKSQKSGVKSTHQRAIVSKIDSRLEQY TSAIEGTAKPTKPAASDLVPVPAEVRNLIKSMWEKGNVFSPTAAGTIPNKETAGLUK GVSRINELWTKTDPGKNSPAPKPSDLRPGDVSSKRNLMWEKQSDVKVTSPTKV" join(230. .850,1616. .2611) /gene="caldesmon" /note="non-muscle form" /codon_start=1 /product="caldesmon" /protein_id="AAA58420.1" /db_xref="GI:180194" /translation="MDDFERRELRRLRKRKEMRLAEIRAYORNDDEEEAARERRRR ARERLRKQEEESLGOVTDQVEVNAQNSVPDEEAKTTTNTQVGDDEAFLERLAR RERERLRQLERLROKQEPDPTITDASLSLPSRMNDTAENETTEKEKSSRORY EIESETVTKSYQXNDWRDASENKEDKEESEEKPKRGSIGENQIKDEKIKKQKE PKBEVKFMDRKKQFTVEKSONGFETHKLKHTENTFRPGGRASVDVTEAEGAPOVE AGKLELRERRGTESEFEKLLKQKQAALELELKKRERRRKVLSEERQRKQE EADRLKEERKRLKEEIERRRAEAERKQWPEGLSDDKKPKFTPKGSSLKIE ERAEFLNKSQKSGVTSQRAIVSKIDSRLEQYTSIAIEGTAKPTKPAASDLVPV AEGVRNLIKSMWEKGNVFSPTAAGTIPNKETAGLUKGVSRINELWTKTDPGKNSPAPK PSDLRPGDVSSKRNLMWEKQSDVKVTSPTKV" 851. .1615 /gene="caldesmon" /note="non-muscle form" 3048 /gene="caldesmon" 3056 /gene="caldesmon" 3124 /gene="caldesmon" 3410 /gene="caldesmon" 3610 /gene="caldesmon"
polyA_site	
polyA_site	
polyA_site	
polyA_site	
polyA_site	
intron	
polyA_site	
polyA_site	
polyA_site	
polyA_site	
polyA_site	
ORIGIN	
Query Match	84.2%; Score 3116.4; DB 9; Length 3610;
Best local Similarity	99.3%; Pred. No. 0;
Matches 3140; Conservative	0; Mismatches 21; Indels 1; Gaps 1;
Qy	217 CCTGCTGCCCGCTGCTTGTCTCTGGGTGGTGGTCTCTTAAAGAAATCAGTCTCTCC 276
Db	1 CTTGCTGCCCGCTGCTTGTCTCTGGGTGGTGGTCTCTTAAAGAAATCAGTCTCTCC 60
Qy	277 TTTCGCACTTAGTCTCCGGGAAGAAGTTTCAGACTACAAGGATATCATTTGGAACATTTCAA 336
Db	61 TTTCGCACTTAGTCTCCGGGAAGAAGTTTCAGACTACAAGGATATCATTTGGAACATTTCAA 120

NOTES

1141	Db	AGAAAGGAAGGAGAGAGGCGAGAGAGAGAGGGAAGGATAGGGAGGAAGAGAAAAGGCG	1200
1417	Qy	AGCAGAGGAGAGGCGAGAGGATAAAGGAGGAGAGAGAAAAGGCGACAGCAGAGAGGCGAGAG	1476
1201	Db	AGCAGAGGAGAGGCGAGAGGATAAAGGAGGAGAGAGAAAAGGCGACAGCAGAGAGGCGAGAG	1260
1477	Qy	GATTAAGGAGGAAGAGAGAAAAGGCGACAGAGGAGAGGCGACAGAGNTAAAGAGAGGAAGAA	1536
1261	Db	GATTAAGGAGGAAGAGAGAAAAGGCGACAGAGGAGAGAGGCGACAGAGNTAAAGAGAGGAAGAA	1320
1537	Qy	AAGGGCAGCAGAGAGGAGGCAAAAGGGCCAGGCGCAGAGGAGAGAGAAAGGCTTAAGGTAGA	1596
1321	Db	AAGGGCAGCAGAGAGGAGGCAAAAGGGCCAGGCGCAGAGGAGAGAGAAAGGCTTAAGGTAGA	1380
1597	Qy	AGAGCGAARACGTAAACAGCAGCTAGBAGAGAGAAARACATGCCATGCAAGAGCAAAAGT	1656
1381	Db	AGAGCGAARACGTAAACAGCAGCTAGBAGAGAGAAARACATGCCATGCAAGAGCAAAAGT	1440
1657	Qy	AAAAGGGGAAAAAGGTAGAAACAGAAAAATAGAAAGGGAATGGGTAAATGAAAAGAAAGCACCA	1716
1441	Db	AAAAGGGGAAAAAGGTAGAAACAGAAAAATAGAAAGGGAATGGGTAAATGAAAAGAAAGCACCA	1500
1717	Qy	AGAAGATAAACTTCAGACAGCTGCTCTAAAGAAACAGGGAGAGAGAGAGGGAACCTAAAGT	1776
1501	Db	AGAAGATAAACTTCAGACAGCTGCTCTAAAGAAACAGGGAGAGAGAGAGGGAACCTAAAGT	1560
1777	Qy	GCAAGCTAAAGAGAAAAAGTTCGAAGAGACAAGCTTACTTCAA AAAAGAGAGATCAA	1836
1561	Db	GCAAGCTAAAGAGAAAAAGTTCGAAGAGACAAGCTTACTTCAA AAAAGAGAGATCAA	1620
1837	Qy	AGATGAAAAGATTAAAAGGACAAAGAACCCAAAGAGAAGTTTAAGAGCTTCATGGATCG	1896
1621	Db	AGATGAAAAGATTAAAAGGACAAAGAACCCAAAGAGAAGTTTAAGAGCTTCATGGATCG	1680
1897	Qy	AAAGAGGGATTTACAGAAGTTAAGTCGAGATPGAGAAATTCAATGACCACAAAACCTTAA	1956
1681	Db	AAAGAGGGATTTACAGAAGTTAAGTCGAGATPGAGAAATTCAATGACCACAAAACCTTAA	1740
1957	Qy	ACATACTGAGAAATACTTTACAGCCCTCGAGGGAGGGCCAGCGTGACACCAAGGAGGC	2016
1741	Db	ACATACTGAGAAATACTTTACAGCCCTCGAGGGAGGGCCAGCGTGACACCAAGGAGGC	1800
2017	Qy	TGAGGGCGCCCCCAGGTGGAAGCGCGCAAAAGGCTGGAGAGCTTCGTGCTGCTCGCGG	2076
1801	Db	TGAGGGCGCCCCCAGGTGGAAGCGCGCAAAAGGCTGGAGAGCTTCGTGCTGCTCGCGG	1860
2077	Qy	GGAGCCGAGAGCGAAGAGTTTCGAGAAGCTCAAAACAGACAGCAGAGAGCGGCTTTGGA	2136
1861	Db	GGAGCCGAGAGCGAAGAGTTTCGAGAAGCTCAAAACAGACAGCAGAGAGCGGCTTTGGA	1920
2137	Qy	GCTGGAGGAATCAAGAAAAAAGAGGGAGGAGAGAAAGGTCTTGAGAGAGGAAGCA	2196
1921	Db	GCTGGAGGAATCAAGAAAAAAGAGGGAGGAGAGAAAGGTCTTGAGAGAGGAAGCA	1980
2197	Qy	GAGGAGGAACGACAGAGGAAGCCGATCGAAAACCTCAGAGAGAGGAAGCAAGAGAGAGGCT	2256
1981	Db	GAGGAGGAACGACAGAGGAAGCCGATCGAAAACCTCAGAGAGAGGAAGCAAGAGAGAGGCT	2040
2257	Qy	AAAGGAGAGATTTCAAAGGGCGAAGCAGAGAGCTGCTGAGAAAGCCAGAGATGCCAGA	2316
2041	Db	AAAGGAGAGATTTCAAAGGGCGAAGCAGAGAGCTGCTGAGAAAGCCAGAGATGCCAGA	2100
2317	Qy	AGATGGCTTGTGATGACAGAAACCATTAAGTGTTCACCTCTAAAGGTTTCACTCT	2376
2101	Db	AGATGGCTTGTGATGACAGAAACCATTAAGTGTTCACCTCTAAAGGTTTCACTCT	2160
2377	Qy	CAAGATAGAGAGCGACAGNAATTTTGAATAGTCTGTCCAGAAAAGCAGTGGTCTCAA	2436
2161	Db	CAAGATAGAGAGCGACAGNAATTTTGAATAGTCTGTCCAGAAAAGCAGTGGTCTCAA	2220
2437	Qy	ATCGACCCATCAAGCAGCAATAGTCTCCAAAGATTGACAGCAGACTGAGCAGATATACCAG	2496

Db	2221	ATCGACCCATCAAGCAGCAATAGTCTCCAAGATTGACAGCAGACTGGAGCAGTATACCAG	2288
Qy	2497	TGCAATTGAGGAAACAAAAGCGCAAAACCTACAAAGCCGCGAGCTCCGGATCTTCCTGT	2556
Db	2281	TGCAATTGAGGAAACAAAAGCGCAAAACCTACAAAGCCGCGAGCTCCGGATCTTCCTGT	2340
Qy	2557	TCCTGCTGAAGTGTACGCAACATCAAGAGTATGCTGGAGAAAGGGAATGCTGTTTCATC	2616
Db	2341	TCCTGCTGAAGTGTACGCAACATCAAGAGTATGCTGGAGAAAGGGAATGCTGTTTCATC	2400
Qy	2617	CCCACTGCGCAGGCACACCAAAATGAAGAAACCTGCTGGCTTTGAAGGTAGGGGTTTCTAG	2676
Db	2401	CCCACTGCGCAGGCACACCAAAATGAAGAAACCTGCTGGCTTTGAAGGTAGGGGTTTCTAG	2460
Qy	2677	CGCATCAATGAATGGCTAACTAAAAACCCAGATGGAACCAAGTCACCTGCTCCCAAAACC	2736
Db	2461	CGCATCAATGAATGGCTAACTAAAAACCCAGATGGAACCAAGTCACCTGCTCCCAAAACC	2520
Qy	2737	TTCTGACTTGACACGAGGAGAGCTATCCAGCAAGCGGAACCTCTGGGAAAGCAATCTGT	2796
Db	2521	TTCTGACTTGACACGAGGAGAGCTATCCAGCAAGCGGAACCTCTGGGAAAGCAATCTGT	2580
Qy	2797	GGATAAGGTCACTTCCCCCACTAAAGGTTTGAGACAGTTCCAGAAAGAACCCCAAGCTCAAG	2856
Db	2581	GGATAAGGTCACTTCCCCCACTAAAGGTTTGAGACAGTTCCAGAAAGAACCCCAAGCTCAAG	2640
Qy	2857	ACGCAAGGACGAGCTCAGTTGTAGAGGCTAACTCGCTCTGTTTTCATATTATGTTCATTT	2916
Db	2641	ACGCAAGGACGAGCTCAGTTGTAGAGGCTAACTCGCTCTGTTTTCATATTATGTTCATTT	2700
Qy	2917	ACTAAATTGGGTTCAATATCTTTTATTTTTCATATATCCAGTAACCCATGTATATATC	2976
Db	2701	ACTAAATTGGGTTCAATATCTTTTATTTTTCATATATCCAGTAACCCATGTATATATC	2760
Qy	2977	ACTATATTATTAATCACACAGTCTAGAGATGTTTCATGTTAAAAGTACTGCCTTTGCACAGG	3036
Db	2761	ACTATATTATTAATCACACAGTCTAGAGATGTTTCATGTTAAAAGTACTGCCTTTGCACAGG	2820
Qy	3037	AGCCTGTTTCTTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCATAAAC	3096
Db	2821	ATCCTGTTTCTTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCATAAAC	2880
Qy	3097	TCTGTATCTGAGCAGTGATACCAACCAACATCTGGAAGTCAACAGAGAATCCAAGTTTAAAA	3156
Db	2881	TCTGTATCTGAGCAGTGATACCAACCAACATCTGGAAGTCAACAGAGAATCCAAGTTTAAAA	2940
Qy	3157	TTGCCCTGGGAATGCTGCAGTATCTGAAAAATGAAACCGTAGTTTTCGTTTTTTTAAAT	3216
Db	2941	TTG-CTGCGGAATGCTGCAGTATCTGAAAAATGAAACCGTAGTTTTCGTTTTTTTAAAT	2999
Qy	3217	ACAGAAGTCATGTGTTTCTGCACTTTTATAATAAGCATGGAAGAAATTAATCTTAGTAGG	3276
Db	3000	ACAGAAGTCATGTGTTTCTGCACTTTTATAATAAGCATGGAAGAAATTAATCTTAGTAGG	3059
Qy	3277	CAATTGTACACTTTTTTGAAGTAAACCCATTTCCAGATTTGAAATACTGCGATAATGGTTG	3336
Db	3060	CAATTGTACACTTTTTTGAAGTAAACCCATTTCCAGATTTGAAATACTGCAATAATGGTTG	3119
Qy	3337	CTTTAAAAAAGAAAAAGATCGCGCAGCAGTTATTCTCTTT	3378
Db	3120	TCTTTAAAAAAGAAAAAGATGTACTGTTAAGGTATTACTTTT	3161

[illegible]

Db 1565 ATCCCCACTGCGAGCGCACACCAATAAGAAATGCTGCTGTAAGGTAGGGTTTC 1624
QY 2674 TAGCCGCATCAATGAATGGCTAACTAAACCCAGATGGAACAAGTCACTGCTCCCAA 2733
Db 1625 TAGCCGCATCAATGAATGGCTAACTAAACCCAGATGGAACAAGTCACTGCTCCCAA 1684
QY 2734 ACCTTCTGACTGAGACGAGAGCGTATCCAGAACGGGAACCTCTGGGAAAAGCAATC 2793
Db 1685 ACCTTCTGACTGAGACGAGAGCGTATCCAGAACGGGAACCTCTGGGAAAAGCAATC 1744
QY 2794 TGTGGATAAGGTCACTTCCCCCACTAAGGTTTGAGACAGATTCCAGAAAAGAACCCCAAGCTC 2853
Db 1745 TGTGGATAAGGTCACTTCCCCCACTAAGGTTTGAGACAGATTCCAGAAAAGAACCCCAAGCTC 1804
QY 2854 AAGACCGAGACAGCTCAGTGTAGAGGGTAATTCGCTCTGTTTGTATTTATTTGA 2913
Db 1805 AAGACCGAGACAGCTCAGTGTAGAGGGTAATTCGCTCTGTTTGTATTTATTTGA 1864
QY 2914 TTACTTAAATTTGGGTTTCATTTATCTTTTATTTTCAATATCCAGTAAACCCATGTATTT 2973
Db 1865 TTACTTAAATTTGGGTTTCATTTATCTTTTATTTTCAATATCCAGTAAACCCATGTATTT 1924
QY 2974 ATCACTATATTTAATATCAAGTCTAGAGTGTTCATGGTAAAGTACTGCTTTGCAC 3033
Db 1925 ATCACTATATTTAATATCAAGTCTAGAGTGTTCATGGTAAAGTACTGCTTTGCAC 1984
QY 3034 AGGAGCTGTCTTAAAGAAACCCATGCTGTGAATATAGAGACTTTTCTACTGATCATCAT 3093
Db 1985 AGGAGCTGTCTTAAAGAAACCCATGCTGTGAATATAGAGACTTTTCTACTGATCATCAT 2044
QY 3094 AACTCTGTATCTGAGCGATGATACCAACCATCTGAAGTCAACAGAAGTCCAAAGTTTA 3153
Db 2045 AACTCTGTATCTGAGCGATGATACCAACCATCTGAAGTCAACAGAAGTCCAAAGTTTA 2104
QY 3154 AAATTCCTCGCGAATGTGTGCAATCTAGAAAATGAACCGTAGTTTGTGTTTTTTA 3213
Db 2105 AAATTCCTCGCGAATGTGTGCAATCTAGAAAATGAACCGTAGTTTGTGTTTTTTA 2164
QY 3214 AATACAGAGTCAATGTTGTTTCTGCATCTTATAATAAGCATTGGAGAAATATCTTAGT 3273
Db 2165 AATACAGAGTCAATGTTGTTTCTGCATCTTATAATAAGCATTGGAGAAATATCTTAGT 2224
QY 3274 AGCAATTTGAACACTTTTGAAGTAAACCATTTTCAGATTGAAATCTGCGATAATGG 3333
Db 2225 AGCAATTTGAACACTTTTGAAGTAAACCATTTTCAGATTGAAATCTGCGATAATGG 2284
QY 3334 TTGCTTTAAAAAAGAAAGATGCGGCGACAGTTATTCCTTT 3378
Db 2285 TTGCTTTAAAAAAGAAAGATGTAAGTGTGTTTAAAGGTATTACTTT 2329

RESULT 7
AR001263 LOCUS 2215 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 14 from patent US 5739008.
ACCESSION AR001263
VERSION AR001263.1 GI:3963330
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2215)
AUTHORS Hayashi,K., Hashida,T., Asada,K., Kotani,H., Kato,I. and Sobue,K.
TITLE DNA encoding a protein comprising calmodulin- and actin-binding
human caldesmon peptide fragment
JOURNAL Patent: US 5739008-A 14 14-APR-1998;
FEATURES
source Location/Qualifiers
1..2215
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 42.4%; Score 1570.2; DB 6; Length 2215;

Best Local Similarity 94.5%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 93; Indels 4; Gaps 3;
QY 1601 CAGAAAACGTAAACAGCAGCTAGAGAGAAAACATGCCATGCAAGAGACAAGATATAAA 1660
Db 462 CAAGAAAGATACAGATAGAGAAACAGAAACAGTCCACCAAGTCTCTACCAAGAAATGAT 521
QY 1661 GGGGAAAAGGTGAACACAGAAAATAGAAAGGGAAATGGGTAAATGAAAGAAAGACACAGAA 1720
Db 522 TGGAGGGATGCTGAAGAAAACAGAAAAGAGACAAGGAAAAGGAGGAGGAGGAGAG 581
QY 1721 GATAAACTTCAGACAGCTGTCTTAAAGAAAACAGGAGAGAGAAAGGAACTAAAGTGCAA 1780
Db 582 AAGCCAAAGCGAGGAGCATTTGGAGAAAATCAGGGAGAGAGAGGAACTAAAGTGCAA 641
QY 1781 GCTAAAGAGAAAAGCTCCAAAGAAAGCAAGCTTACCTTCAAAAAGAGAGATCAAGAT 1840
Db 642 GCTAAAGAGAAAAGCTCCAAAGAAAGCAAGCTTACCTTCAAAAAGAGAGATCAAGAT 701
QY 1841 GAAAAGATTAAAAAGGACAAAGAACCCAAAGAAAGTTAAGAGCTTCATGGATCGAAAG 1900
Db 702 GAAAAGATTAAAAAGGACAAAGAACCCAAAGAAAGTTAAGAGCTTCATGGATCGAAAG 761
QY 1901 AAGGGATTTACAGAGTTAAGTCCAGAAATGAGAAATTCATGACCCACAACTTAAACAT 1960
Db 762 AAGGGATTTACAGAGTTAAGTCCAGAAATGAGAAATTCATGACCCACAACTTAAACAT 821
QY 1961 ACTGAGAAATACTTTACGCCCTTGGAGGGAGGGCCAGCTGGACACCAAGAGAGCTGAG 2020
Db 822 ACTGAGAAATACTTTACGCCCTTGGAGGGAGGGCCAGCTGGACACCAAGAGAGCTGAG 881
QY 2021 GGGCCCCCAGGTGGAAGCGGCAAAAAGGTGAGAGAGCTTCGTCTGCTCGCGGGAG 2080
Db 882 GGGCCCCCAGGTGGAAGCGGCAAAAAGGTGAGAGAGCTTCGTCTGCTCGCGGGAG 941
QY 2081 ACCGAGAGCGAAGTTCGAGAACTCAACAGAAAGCAGCAGAGGCGGCTTTGAGAGCTG 2140
Db 942 ACCGAGAGCGAAGTTCGAGAACTCAACAGAAAGCAGCAGAGGCGGCTTTGAGAGCTG 1001
QY 2141 GAGGAATCTAAGAAAAGAGGGAGAGAGAAAGGTCTGGAGAGGAGAAAGAGAGAGG 2200
Db 1002 GAGGAATCTAAGAAAAGAGGGAGAGAGAAAGGTCTGGAGAGGAGAAAGAGAGAGG 1061
QY 2201 AGGAAGCAGAGAGAGCGGATCGAAAATCTAGAGAGGAGGAGAGAGAGAGGCTTAAAG 2260
Db 1062 AGGAAGCAGAGAGAGCGGATCGAAAATCTAGAGAGGAGGAGAGAGAGAGGCTTAAAG 1121
QY 2261 GAAGAGATTGAAAGCGGAGAGCAGAGAGCTGCTGAGAAAACGCCAGAAAGTCCAGAGAT 2320
Db 1122 GAAGAGATTGAAAGCGGAGAGCAGAGAGCTGCTGAGAAAACGCCAGAAAGTCCAGAGAT 1181
QY 2321 GGCTTGTCAATGACAAAGAACCAATTCAGAGTGTCTTCACTCTTAAAGGTTCACTCTCAAG 2380
Db 1182 GGCTTGTCAATGACAAAGAACCAATTCAGAGTGTCTTCACTCTTAAAGGTTCACTCTCAAG 1241
QY 2381 ATAGAGAGCGAGCAGAAATTTTGAATAAGTCTGTGAGAAAAGAGCAGTGGTGTCAAATCG 2440
Db 1242 ATAGAGAGCGAGCAGAAATTTTGAATAAGTCTGTGAGAAAAGAGCAGTGGTGTCAAATCG 1301
QY 2441 ACCCATCAAGCAGCAATAGTCTCCCAAGATTGACAGCAGCTGGAGCAGTATACCACTGCA 2500
Db 1302 ACCCATCAAGCAGCAATAGTCTCCCAAGATTGACAGCAGCTGGAGCAGTATACCACTGCA 1361
QY 2501 ATTGAGGAAACAAAAGCGCAAAACCTTACAAAGCGGAGGCTCGGATCTTCTGTCTCT 2560
Db 1362 ATTGAGGAAACAAAAGCGCAAAACCTTACAAAGCGGAGGCTCGGATCTTCTGTCTCT 1421
QY 2561 GCTGAGAGTGTACGCAACATCAAGATATGTGGGAGAAAGGGAATGTGTTTTCATCCCC 2620
Db 1422 GCTGAGAGTGTACGCAACATCAAGATATGTGGGAGAAAGGGAATGTGTTTTCATCCCC 1481
QY 2621 ACTGCAAGAGGACACCAAAATAAGAAAATGCTGCTGCTTAAAGGTAGGGGTCTTACCGCG 2680

1482 ACTGACGAGCAGCACCAAAATAAGGAACTGCTGGCTTGAAGTAGGGGTTTCTAGCCGT 1541
2681 ATCAATGAATGGTAACTAAACCCAGATGGAACAAAGTCACCTGCTCCCAAACTTCT 2740
1542 ATCAATGAATGGTAACTAAACCCAGATGGAACAAAGTCACCTGCTCCCAAACTTCT 1601
2741 GACTTGAGACAGGAGAGCTATCCAGCAAGCGAACCTCTGGGAAAGCAATCTGTGGAT 2800
1602 GACTTGAGACAGGAGAGCTATCCAGCAAGCGAACCTCTGGGAAAGCAATCTGTGGAT 1661
2801 AAGTCTACTTCCCACTAAGTGTGAGACAGTTCAGAAAGAACCCAAAGCTCAAGACGC 2860
1662 AAGTCTACTTCCCACTAAGTGTGAGACAGTTCAGAAAGAACCCAAAGCTCAAGACGC 1721
2861 AGGACGAGCTCAGTGTAGAGGGCTAAATCGCTGCTGTTTGTATTTATGTTTACTA 2920
1722 AGGACGAGCTCAGTGTAGAGGGCTAAATCGCTGCTGTTTGTATTTATGTTTACTA 1779
2921 AATTGGGTTCAATTATCTTTTATTTTCAATATCCAGTAAACCCATGTATATTCACTA 2980
1780 AATTGGGTTCAATTATCTTTTATTTTCAATATCCAGTAAACCCATGTATATTCACTA 1839
2981 TATTATAATACAGTCTAGAGATGTTTCATGTTAAAGTACTGCTTTTGCACAGGAGCC 3040
1840 TATTATAATACAG-CTAGAGATGTTTCATGTTAAAGTACTGCTTTTGCACAGGAGCC 1898
3041 TGTTTCTAAAGAAACCCATGCTGTGAAATPAGAGACTTTTCTACTGATCATATAAATCTG 3100
1899 TGTTTCTAAAGAAACCCATGCTGTGAAATPAGAGACTTTTCTACTGATCATATAAATCTG 1958
3101 TATCTGAGCAGTATACCAACCATCTGAAGTCAACAGAGATCCAAAGTTTAAATATGC 3160
1959 TATCTGAGCAGTATACCAACCATCTGAAGTCAACAGAGATCCAAAGTTTAAATATGC 2018
3161 CTGCGGAATGTGTCAGTATCTAGAAAATGAACCGTAGTTTGTGTTTAAATACAG 3220
2019 CTGCGGAATGTGTCAGTATCTAGAAAATGAACCGTAG-TTTTGTGTTTAAATACAG 2077
3221 AAGTCATGTTGTTCTGCACTTTATAATAAGCATGGAAGAAATTTATCTAGTAGCAAT 3280
2078 AAGTCATGTTGTTCTGCACTTTATAATAAGCATGGAAGAAATTTATCTAGTAGCAAT 2137
3281 TGTAACTCTTTTGAAGTAAACCCATTTAGAAATTTGAAATCTGCGATATGTTGTTT 3340
2138 TGTAACTCTTTTGAAGTAAACCCATTTAGAAATTTGAAATCTGCGATATGTTGTTT 2197
3341 AAAAAAAAAAAGAAAGA 3357
2198 TAAAAAAAAAAAAAAAAA 2214

RESULT 8
E05382
LOCUS
DEFINITION DNA encoding human caldesmon.
ACCESSION E05382
VERSION E05382.1 GI:2173571
KEYWORDS JP 1993230099-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2215)
Hayashi,K., Hashida,H., Asada,K., Kotani,H., Katou,I. and Sofue,K.
POLYPEPTIDE
Patent: JP 1993230099-A 3 07-SEP-1993;
SOFUE KENJI, TAKARA SHUZO CO LTD
OS Homo sapiens (human)
PN JP 1993230099-A/3
PD 07-SEP-1993
PF 27-DEC-1991 JP 1991358040
PR 29-MAR-1991 JP 91P 89106
PI HAYASHI KENICHIRO, HASHIDA HISASHI, ASADA KIYOZOU, PI KOTANI

HIROICHI,
PI KATOU IKUNOSHIN, SOFUE KENJI
PC C07K13/00,C12N15/12//A61K37/02,C12P21/02,G01N33/53,(C12P21/02,
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC *source: cell_type=Hela; Location/Qualifiers
FH Key
FH CDS
FT 12..1688
FT /product='human caldesmon'
FT polyA signal 2103..2108
FT polyA site 2199..2215.
FT Location/Qualifiers
FEATURES
source
1..2215
/organism='Homo sapiens'
/mol_type='genomic RNA'
/db_xref='taxon:9606'

ORIGIN
Query Match 42.4%; Score 1570.2; DB 6; Length 2215;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 93; Indels 4; Gaps 3;
Qy 1601 CAGAAACCTAACAGCAGCTAGAGAGAAACATGCTCCATGCAAGAGACAAAGATAAAA 1660
Db 462 CAAGAAAGATACGAGATAGAGAAACAGAAACAGTCACCAAGTCCTACCAAGAGATGAT 521
Qy 1661 GGGGAAAGGTAGAACAGAAATAGAGGGGAATGGTAAATGAAAAAGAAAGCAAGAA 1720
Db 522 TGGAGGATGTCTAAGAAACCAAGAAAGAAAGAGGAGGAGGAGAGAGAG 581
Qy 1721 GATAAATTCAGACAGCTGTCTTAAAGAAACAGGAGAGAGAGAGGAACTTAAAGTGCAA 1780
Db 582 AAGCAAAAGCGAGGAGCATTTGGAGAAATTCAGGAGAGAGAGAGGAACTTAAAGTGCAA 641
Qy 1781 GCTAAACAGAGAAAGCTCCAGAGACAGACCTTACCTTCAAAAAAGAGAGATCAAGAT 1840
Db 642 GCTAAAGAGAAAGCTCCAGAGAGACAGCTTACCTTCAAAAAAGAGAGATCAAGAT 701
Qy 1841 GAAAGATTAAAAAGCAGAAAGACCCAAAGAAAGTTAAAGAGCTTCTGATCGAAAG 1900
Db 702 GAAAGATTAAAAAGCAGAAAGACCCAAAGAAAGTTAAAGAGCTTCTGATCGAAAG 761
Qy 1901 AAGGATTTACAGAGTTAAGTCGAGATGAGATTCATGACCCCAACATTTAAACAT 1960
Db 762 AAGGATTTACAGAGTTAAGTCGAGATGAGATTCATGACCCCAACATTTAAACAT 821
Qy 1961 ACTGAGAAATCTTTTCAGCGCCCTTGAGGAGGCGGCGTGGACACCAAGAGGCTGAG 2020
Db 822 ACTGAGAAATCTTTTCAGCGCCCTTGAGGAGGCGGCGTGGACACCAAGAGGCTGAG 881
Qy 2021 GCGCCCCCAGGTGGAAGCCGCGCAAGGCTGGAGAGGTTCTGTCGTCGCGGGGAG 2080
Db 882 GCGCCCCCAGGTGGAAGCCGCGCAAGGCTGGAGAGGTTCTGTCGTCGCGGGGAG 941
Qy 2081 ACCGAGAGCAGAGAGCTTCGAGAAAGCTCAACAGAGCAGCAGGAGGCGGCTTTGGAGCTG 2140
Db 942 ACCGAGAGCAGAGAGCTTCGAGAAAGCTCAACAGAGCAGCAGGAGGCGGCTTTGGAGCTG 1001
Qy 2141 GAGGACTCAGAAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2200
Db 1002 GAGGAACTCAAGAAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1061
Qy 2201 AGGAGAGGAG 2260
Db 1062 AGGAGAGGAG 1121
Qy 2261 GAGAGAGTTGAAAGCGAG 2320
Db 1122 GAGAGAGTTGAAAGCGAG 1181
Qy 2321 GGCTGTGATGACAG 2380

1152 GCCTTGTCAGATGACAGAAACCACTCAAGTGTTCCTCTAAAGGTTCTATCTCTCAAG 1241
 2381 ATAGAAGAGCGAGCAGCAATTTTGAATGAATGCTGTGCAGAAAGCAGTGGTGTCAATCG 2440
 1242 ATAGAAGAGCGAGCAGCAATTTTGAATGAATGCTGTGCAGAAAGCAGTGGTGTCAATCG 1301
 2441 ACCCATCAAGCAGCAATAGTCTCCAGATTGACAGCAGACTGGAGCAGATACCAAGTCCA 2500
 1302 ACCCATCAAGCAGCAATAGTCTCCAGATTGACAGCAGACTGGAGCAGATACCAAGTCCA 1361
 2501 ATTGAGGGAACAAAGAGCGCAAACTTACAAAGCGCGCAGCTCGATCTTCTGTTCTCT 2560
 1362 ATTGAGGGAACAAAGAGCGCAAACTTACAAAGCGCGCAGCTCGATCTTCTGTTCTCT 1421
 2561 GCTGAAGTGTACGCAATCAAGAGTATGTGGAGAAAGGAATGTGTTTTCATCCCCC 2620
 1422 GCTGAAGTGTACGCAATCAAGAGTATGTGGAGAAAGGAATGTGTTTTCATCCCCC 1481
 2621 ACTGAGCAGGACACCAATATAGGAACCTGCTGGCTTGAAGTAGGGGTTTCTAGCCGC 2680
 1482 ACTGAGCAGGACACCAATATAGGAACCTGCTGGCTTGAAGTAGGGGTTTCTAGCCGC 1541
 2681 ATCAATGAATGGCTAACTAAACCCCAAGTGAAGTCAAGTCAAGTCTCCCAACCTTCT 2740
 1542 ATCAATGAATGGCTAACTAAACCCCAAGTGAAGTCAAGTCAAGTCTCCCAACCTTCT 1601
 2741 GACTTGAGACCGAGCAGTATCCAGAGCGGAACCTCTGGAAAGCAATCTGTGGAT 2800
 1602 GACTTGAGACCGAGCAGTATCCAGAGCGGAACCTCTGGAAAGCAATCTGTGGAT 1661
 2801 AAGGTCACCTTCCCCTACTAAGGTTTGAGCAGTTCCAGAAAGAACCCCAAGTCTCAAGACG 2860
 1662 AAGGTCACCTTCCCCTACTAAGGTTTGAGCAGTTCCAGAAAGAACCCCAAGTCTCAAGACG 1721
 2861 AGGACAGCTCAGTTGTAGAGGCTATTCCTGCTGTTGTTATTTATTTATTTATTTATTT 2920
 1722 AGGACAGCTCAGTTGTAGAGGCTATTCCTGCTGTTGTTATTTATTTATTTATTTATTT 1779
 2921 AATTGGGTTTCATTATCTTTTATTTTCAATATCCAGTAAACCCATGATATTTATTTATTT 2980
 1780 AATTGGGTTTCATTATCTTTTATTTTCAATATCCAGTAAACCCATGATATTTATTTATTT 1839
 2981 TATTTAATATCAAGTCTAGAGATGTTCTATGTTAAAGTACTGCTTTTGACAGGAGCC 3040
 1840 TATTTAATATCAAG-CTAGAGATGTTCTATGTTAAAGTACTGCTTTTGACAGGAGCC 1898
 3041 TGTTCCTAAAGAAACCCATGCTGTGAATAGAGACTTTTCTACTGATCATCTAATCTCTG 3100
 1899 TGTTCCTAAAGAAACCCATGCTGTGAATAGAGACTTTTCTACTGATCATCTAATCTCTG 1958
 3101 TATCTGAGCAGTATACCAACCCATCTGAAGTCAACAGAGATCCAAAGTTTAAATTCG 3160
 1959 TATCTGAGCAGTATACCAACCCATCTGAAGTCAACAGAGATCCAAAGTTTAAATTCG 2018
 3161 CTGCGGAATGTGTGAGTATCTAGAAATTAAGACCGTAGTTTGTGTTTAAATACAG 3220
 2019 CTGCGGAATGTGTGAGTATCTAGAAATTAAGACCGTAG-TTTTGTTTTTAAATACAG 2077
 3221 AAGTCATGTTGTTTCTGCACTTTATATAAAGCATGGAAGAAATTAATCTAGTAGGCAAT 3280
 2078 AAGTCATGTTGTTTCTGCACTTTATATAAAGCATGGAAGAAATTAATCTAGTAGGCAAT 2137
 3281 TGTAAACACTTTTGAAGTAACCATTTTCAGATTGAAATCTCGATAATGTTGTTT 3340
 2138 TGTAAACACTTTTGAAGTAACCATTTTCAGATTGAAATCTCGATAATGTTGTTT 2197
 3341 AAAAAAAAAAAAAA 3357
 2198 TAAAAAAAAAAAAA 2214

RESULT 9

123421
 LOCUS 2215 bp DNA linear PAT 07-OCT-1996
 DEFINITION Sequence 14 from patent US 5532337.
 ACCESSION 123421
 VERSION 123421.1 GI:1603291
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2215)
 AUTHORS Hayashi,K., Hashida,T., Asada,K., Kotani,H., Kato,I. and Sobue,K.
 TITLE Proteins comprising calmodulin- and actin-binding human caldesmon
 JOURNAL Polyptide fragments
 FEATURES Patent: US 5532337-A 14 02-JUL-1996;
 Location/Qualifiers
 1. .2215
 /organism="unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Query Match 42.4%; Score 1570.2; DB 6; Length 2215;
 Best Local Similarity 94.5%; Pred. No. 0; Mismatches 93; Indels 4; Gaps 3;
 Matches 1660; Conservative 0;
 QY 1601 CAGAAACGTAACAGCAGCTAGAGAGAAACCAATGCCATGCCATGCCAGACAAAGATAAAA 1660
 DB 462 CAAGAAAGATACGAGATAGAGGAAACAGAAACAGTCCCAAGTCCCTACAGAAAGATGAT 521
 QY 1661 GGGGAAAGGTAGACAGAAATAGAGGGAATGGTAAATGAAGAAAGACACAAAGAA 1720
 DB 522 TGGAGGATGCTGAAGAAACCAAGAAAGACAAAGGAAAGGAGGAGGAGAGAG 581
 QY 1721 GATAAACTTCAGACAGCTGCTTAAAGAAACAGGAGAGAGAGAGGAACTTAAAGTGCAA 1780
 DB 582 AGGCCAAGCGGGAGCATTGGAGAAATCAGGAGAGAGAGAGGGAACCTAAAGTGCAA 641
 QY 1781 GCTAAAGAGAAAGCTCCAGAAAGACAGCTACCTTCAAAAAGAGAGAGATCAAAGAT 1840
 DB 642 GCTAAAGAGAAAGCTCCAGAAAGACAGCTACCTTCAAAAAGAGAGAGATCAAAGAT 701
 QY 1841 GAAAGATTAAAAAGGACAAAGAACCCAAAGAAAGTAAAGAGCTTCATGGATCGAAAG 1900
 DB 702 GAAAGATTAAAGAGCAAGAACCCAAAGAAAGTAAAGAGCTTCATGGATCGAAAG 761
 QY 1901 AAGGATTTACAGAGTTAAGTCCAGAAATGAGAAATTCATGCCACCAAACTTAAACAT 1960
 DB 762 AAGGATTTACAGAGTTAAGTCCAGAAATGAGAAATTCATGCCACCAAACTTAAACAT 821
 QY 1961 ACTGAGAACTTTTACGCGCCCTTGGAGGAGGCGCAGCTGGACACCAAGAGGCTGAG 2020
 DB 822 ACTGAGAACTTTTACGCGCCCTTGGAGGAGGCGCAGCTGGACACCAAGAGGCTGAG 881
 QY 2021 GCGCCCCCAGGTGGAAGCGGCAAAAGGCTGGAGGAGCTTGGTCTGCTCGCGGGAG 2080
 DB 882 GCGCCCCCAGGTGGAAGCGGCAAAAGGCTGGAGGAGCTTGGTCTGCTCGCGGGAG 941
 QY 2081 ACCGAGAGCGAGAGTTTCGAGAGAGCTCAAACAGAAAGCAGCAGGAGGCGGCTTGGAGCTG 2140
 DB 942 ACCGAGAGCGAGAGTTTCGAGAGAGCTCAAACAGAAAGCAGCAGGAGGCGGCTTGGAGCTG 1001
 QY 2141 GAGGAACTCAAGAAAAGAGGAGGAGAGAGAGGAGTCTTGGAGGAGGAGAGAGAGG 2200
 DB 1002 GAGGAACTCAAGAAAAGAGGAGGAGAGAGAGGAGTCTTGGAGGAGGAGAGAGAGG 1061
 QY 2201 AGGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2260
 DB 1062 AGGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1121
 QY 2261 GAAGAGATTGAAAGGCGAGAGAGAGAGAGAGTCTGCTGAGAAACCCAGAGAGATCCAGAGAT 2320
 DB 1122 GAAGAGATTGAAAGGCGAGAGAGAGAGAGTCTGCTGAGAAACCCAGAGAGATCCAGAGAT 1181
 QY 2321 GCCTTGTGATGACAGAAACCAATTCAGTGTGTTTCACTCTCTCTCTCTCTCTCTCTCTCAAG 2380

1182 GCCTTGTGAGATGACAGAAACCATTCAGTGTTCCTCCTTAAGGTTCACTCTCAAG 1241
 2381 ATAGAAGCGAGCAGAAATTTTGAATAAGTCTGTGCAGAAAAGCAGTGTGTCAAAATCG 2440
 1242 ATAGAAGCGAGCAGAAATTTTGAATAAGTCTGTGCAGAAAAGCAGTGTGTCAAAATCG 1301
 2441 ACCCATCAAGCAGCAATAGTCTCCAGATTGACAGCAGAGCTGGAGCAGTATACAGTGCA 2500
 1302 ACCCATCAAGCAGCAATAGTCTCCAGATTGACAGCAGAGCTGGAGCAGTATACAGTGCA 1361
 2501 ATTGAGGGAACAAAAGCGCAAAACCTACAAAGCGCGACCTCGATCTTCTGTTCTCT 2560
 1362 ATTGAGGGAACAAAAGCGCAAAACCTACAAAGCGCGACCTCGATCTTCTGTTCTCT 1421
 2561 GCTGAAGGTGTACGCAACATCAAGAGTATGTGGAGAAAGGGAATGTGTTTATCCTCCC 2620
 1422 GCTGAAGGTGTACGCAACATCAAGAGTATGTGGAGAAAGGGAATGTGTTTATCCTCCC 1481
 2621 ACTGCAGCAGGACACACCAATTAAGGAACTGCTGGCTTGAAGTGGGGTTCCTAGCCGC 2680
 1482 ACTGCAGCAGGACACACCAATTAAGGAACTGCTGGCTTGAAGTGGGGTTCCTAGCCGT 1541
 2681 ATCAATGATGCTTAACAAACCCAGATGGAACAGTCACTGCTCCCAACCTTCT 2740
 1542 ATCAATGATGCTTAACAAACCCAGATGGAACAGTCACTGCTCCCAACCTTCT 1601
 2741 GACTTGACACAGGAGACGTATCCAGCAGCGGAACCTCTGGGAAAGCAATCTGTGGAT 2800
 1602 GACTTGACACAGGAGACGTATCCAGCAGCGGAACCTCTGGGAAAGCAATCTGTGGAT 1661
 2801 AAGTCACTTCCCCCACTAAGTTTGAGACAGTTCAGAAAGACCAAGCTCAAGACGC 2860
 1662 AAGTCACTTCCCCCACTAAGTTTGAGACAGTTCAGAAAGACCAAGCTCAAGACGC 1721
 2861 AGGACGAGCTCAGTTGTAGAGGCTAATTCGCTCTGTTTGTATTTATGTTGTTTACTA 2920
 1722 AGGACGAGCTCAGTTGTAGAGGCTAATTCGCTC--TTTGTATTTATGTTGTTTACTA 1779
 2921 AATTGGGTTCAATATCTTTTATTTTCAATATCCAGTAAACCATGTATATTATCACTA 2980
 1780 AATTGGGTTCAATATCTTTTATTTTCAATATCCAGTAAACCATGTATATTATCACTA 1839
 2981 TATTTAATATCACAGTCTAGAGATGTTTCATGTTAAAGTACTGCTTTCACAGGAGCC 3040
 1840 TATTTAATATCACAG-CTAGAGATGTTTCATGTTAAAGTACTGCTTTCACAGGAGCC 1898
 3041 TGTCTTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCAATCACTCTG 3100
 1899 TGTCTTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCAATCACTCTG 1958
 3101 TATCTGAGCAGTGATACCAACACATCTGAAGTCAACAGAGATCCAAAGTTTAAATTCG 3160
 1959 TATCTGAGCAGTGATACCAACACATCTGAAGTCAACAGAGATCCAAAGTTTAAATTCG 2018
 3161 CTGCGGAATGTGTGAGTATCTAGAAAATGAACCGTAGTTTGTGTTTAAATACAG 3220
 2019 CTGCGGAATGTGTGAGTATCTAGAAAATGAACCGTAG-TTTTGTGTTTAAATACAG 2077
 3221 AAGTCATGTTGTTCTGCACTTTATATAAAGCATGGAAGAAATTAATTCTTAGTAGGCAAT 3280
 2078 AAGTCATGTTGTTCTGCACTTTATATAAAGCATGGAAGAAATTAATTCTTAGTAGGCAAT 2137
 3281 TGTAAACCTTTTGAAGTAACCCATTTTCAGATTTCGAATTCGGAATATGTTGCTTT 3340
 2138 TGTAAACCTTTTGAAGTAACCCATTTTCAGATTTCGAATTCGGAATATGTTGCTTT 2197
 3341 AAAAAAAAAAGAAAAA 3357
 2198 TAAAAAAAAAAAAAAAAA 2214

RESULT 10

BC040354 4130 bp mRNA linear PRI 07-OCT-2003
 LOCUS Homo sapiens caldesmon 1, transcript variant 3, mRNA (cdna clone
 DEFINITION MGC:21352 IMAGE:4753285), complete cds.
 ACCESSION BC040354
 VERSION BC040354.1 GI:25955665
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 4130)
 AUTHORS Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, P., Dackenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, T.E., Brownstein, M.J., Usdin, T.B., Casavani, T.D., Scheetz, M.B., Soares, M.B., Donald, M.F., Casavani, T.D., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worthy, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalski, U., Smalinski, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE 12477932
 PUBMED 12388257
 REFERENCE 2 (bases 1 to 4130)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (27-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
 COMMENT Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
 Series: IRAC Plate: 28 Row: p Column: 21
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 15149468.
 FEATURES
 Location/Qualifiers
 1. 4130
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:21352 IMAGE:4753285"
 /tissue_type="Skin, squamous cell carcinoma"
 /clone_lib="NCI CGAP_Skn4"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6.1"
 1. 4130
 gene

CDS	/gene="CALDI"
	/note="synonyms: CDM, NAG22"
	/db_xref="LocusID:800"
	/db_xref="MIM:114213"
	202_..1818
	/codon_start=1
	/product="caldesmon 1, isoform 3"
	/protein_id="AAH40354.1"
	/db_xref="GI:2595566"
	/db_xref="LocusID:800"
	/translation="MDQFERRRLRQRREMRLEARIAYQRNDDDEEAEARRRRR ARQERLRQBEESLQCTDQVEVNAQNSVPDEAKTTTNTQVGGDDDEAFLERLAR REERQKRQLQEALERQKEFPDTTDASLSPPSRMQNDTAENETTEKESSESORERY ETETEVTYKSYQNDRDAENKEDKEEBEEKPGRGSI GSNQIKDKIKKKDKK PKEEVKSFMDRKXGFTEVSQNGEFMTHLKLHTLFELULKKEPERKVL EEPGRKOE AKRIELEIRRGSETSESEPELKQKOQBALTELKKEERKVL EEPGRKOE EADRLREEEKRLKEBTERRAEAKEKQMPEDGLSDKKPKFCPTTPKGSSLKIE EAEFLNKSVQSGGVKSTHQAIVSKIDSLRLEQTTSIALETGTAKIPKAASDLUPV AEGVINKSMWKEGNVFSPTAAQTPNKETAGLVKGVSSRINEMLTITPDGNKSPAPV PSDLRPDGVDYSKRNLEWEKOSVDKVTSPTKV" 202_..735
misc_feature	/note="Caldesmon"
	/db_xref="CDD:pfam02029"
misc_feature	874_..1038
	/note="Caldesmon"
misc_feature	/db_xref="CDD:pfam02029"
	1294_..1815
	/note="Caldesmon"
	/db_xref="CDD:pfam02029"
ORIGIN	
	Query Match 42.2%; Score 1562; DB 9; Length 4130;
	Best Local Similarity 75.2%; Pred. No. 0;
	Matches 2340; Conservative 0; Mismatches 5; Indels 765; Gaps 1;
Qy	245 CTGTCTCTCTCTTAAGAATAACAGTCCTTCTCTTTCGACTTAGTCCTCGGAAGAAGATT 304
Dd	1 CTGTCTCTCTCTTAAGAATAACAGTCCTTCTCTTTCGACTTAGTCCTCGGAAGAAGATT 60
Qy	305 TCAGACTCAAGATCATTTGGAACATTTCAAGATCATCAATCAATCAAATCCACAGGAT 364
Dd	61 TCAGACTCAAGATCATTTGGAACATTTCAAGATCATCAATCAATCAAATCCACAGGAT 120
Qy	365 GGTGACCAACAGAGCGTCAGACATCTGATTTGCTGACTCTGCAGACATCATCTGGTCT 424
Dd	121 GGTGACCAACAGAGCGTCAGACATCTGATTTGCTGACTCTGCAGACATCATCTGGTCT 180
Qy	425 CCTCAACTGAAAATCACACCATTGATTTTCAGCGCTGCCAGAGAACCTTAGAAGGCAA 484
Dd	181 CCTGAACTGAAATCACACCATTGATTTTCAGCGCTGCCAGAGAACCTTAGAAGGCAA 240
Qy	485 AAGAGGGAGGAGATGGACTCGAAGCAGAAAGATCGCTACAGAGGAATGACGATGAT 544
Dd	241 AAGAGGGAGGAGATGGACTCGAAGCAGAAAGATCGCTACAGAGGAATGACGATGAT 300
Qy	545 GAAGAGGAGGAGCGCCGGGAACCGCGCCGCCAGCCGACAGGAACGGCTGCGGAGGAAG 604
Dd	301 GAAGAGGAGGAGCGCCGGGAACCGCGCCGCCAGCCGACAGGAACGGCTGCGGAGGAAG 360
Qy	605 CAGGAGGAGGAGATCTTGACAGAGTACCGACAGGTGGAGGTGAATGCCAGACAGCT 664
Dd	361 CAGGAGGAGGAGATCTTGAGACAGGTACCGACAGGTGGAGGTGAATGCCAGACAGCT 420
Qy	665 GTCCCTGACGAGGAGGCCAAGACAACACCAACCAACACTCAAGTGGAGGGGATGATGAG 724
Dd	421 GTCCCTGACGAGGAGGCCAAGACAACACCAACCAACACTCAAGTGGAGGGGATGATGAG 480
Qy	725 GCCGATTCCTGGAGCGCTGGCTCGCGGTGAGGAAGACGCCAAGACCGCTTCAGGAG 784
Dd	481 GCCGATTCCTGGAGCGCTGGCTCGCGGTGAGGAAGACGCCAAGACCGCTTCAGGAG 540
Qy	785 GCTCTGGAGCGGAGAGGAGTTCGACCCCAACAATAACAGATGCAAGTCTGTCTGCTCCA 844

2y	1925	CAGAA	TGGAG	AATTC	ATGAC	CCCA	CAAA	CTTAA	ACAT	ACTG	AGAA	TACTTT	CAGC	CGCCT	1984
2b	916	CAGAA	TGGAG	AATTC	ATGAC	CCCA	CAAA	CTTAA	ACAT	ACTG	AGAA	TACTTT	CAGC	CGCCT	975
2y	1985	GGAGG	AGGGC	CAGCG	TGCAC	CAAC	AGGAG	CGCT	TAGG	GGCG	CCCC	CCCG	AGTGG	AGCCGGC	2044
2b	976	GGAGG	AGGGC	CAGCG	TGCAC	CAAC	AGGAG	CGCT	TAGG	GGCG	CCCC	CCCG	AGTGG	AGCCGGC	1035
2y	2045	AAAAG	CTGG	AGGAG	CTTCGT	CGTCT	CGCG	GGAG	ACCG	AGCG	AGGA	AGAGTT	TCGAG	AAG	2104
2b	1036	AAAAG	CTGG	AGGAG	CTTCGT	CGTCT	CGCG	GGAG	ACCG	AGCG	AGGA	AGAGTT	TCGAG	AAG	1095
2y	2105	CTCAAA	CAGAA	GCAG	CAGG	AGCG	CGGT	TTTGG	AGCT	GGAG	GAACT	CAAG	AAAA	AGAGG	2164
2b	1096	CTCAAA	CAGAA	GCAG	CAGG	AGCG	CGGT	TTTGG	AGCT	GGAG	GAACT	CAAG	AAAA	AGAGG	1155
2y	2165	GAGAG	AGGAG	GTCT	TGAG	GAGG	AGAG	CAG	CAG	AGG	AGG	AGAG	CAGG	AGCCG	2224
2b	1156	GAGAG	AGGAG	GTCT	TGAG	GAGG	AGAG	CAG	CAG	AGG	AGG	AGAG	CAGG	AGCCG	1215
2y	2225	AAACT	CAGAG	GGAG	AAGAG	AGGAG	GGCT	TAA	GGA	AGAG	ATG	TGA	AAG	CGGA	2284
2b	1216	AAACT	CAGAG	GGAG	AAGAG	AGGAG	GGCT	TAA	GGA	AGAG	ATG	TGA	AAG	CGGA	1275
2y	2285	GAAGT	CTGT	GAAA	AGCC	CAG	AGAG	TG	CC	AGA	AGAT	CG	CTT	G	2344
2b	1276	GAAGT	CTGT	GAAA	AGCC	CAG	AGAG	TG	CC	AGA	AGAT	CG	CTT	G	1335
2y	2345	TTCAAG	TGTT	CACT	CTCT	TAA	AGG	TTCA	TTCT	CA	AGAT	TAG	AGAG	CGAG	2404
2b	1336	TTCAAG	TGTT	CACT	CTCT	TAA	AGG	TTCA	TTCT	CA	AGAT	TAG	AGAG	CGAG	1395
2y	2405	AATAAG	TCTGT	GCAG	AAAA	GCAG	TGTGT	CAAA	TCG	AC	CCAT	CAAG	CAG	CAAT	2464
2b	1396	AATAAG	TCTGT	GCAG	AAAA	GCAG	TGTGT	CAAA	TCG	AC	CCAT	CAAG	CAG	CAAT	1455
2y	2465	AAGATT	GAC	GAC	GACT	TGAG	CGAG	TATAC	CAG	TGC	CAAT	TG	GGGA	CAAA	2524
2b	1456	AAGATT	GAC	GAC	GACT	TGAG	CGAG	TATAC	CAG	TGC	CAAT	TG	GGGA	CAAA	1515
2y	2525	CCTACAA	AGCC	GGC	AGC	CTCG	GA	TCTT	CTCT	TGTT	CTCT	GTG	AGG	TGTAC	2584
2b	1516	CCTACAA	AGCC	GGC	AGC	CTCG	GA	TCTT	CTCT	TGTT	CTCT	GTG	AGG	TGTAC	1575
2y	2585	AGTAT	GTC	GG	AAA	AGG	GA	TGTG	TTT	TCAT	CCCC	CTG	CAG	CAGG	2644
2b	1576	AGTAT	GTC	GG	AAA	AGG	GA	TGTG	TTT	TCAT	CCCC	CTG	CAG	CAGG	1635
2y	2645	GAAACT	GTC	TGG	CTT	GAAG	TAG	GGG	TTT	CTAG	CG	CA	TCA	TGA	2704
2b	1636	GAAACT	GTC	TGG	CTT	GAAG	TAG	GGG	TTT	CTAG	CG	CA	TCA	TGA	1695
2y	2705	CCAGAT	TG	AAA	CAAG	TCA	CTG	CTC	CA	AA	ACCT	CTG	AC	T	2764
2b	1696	CCAGAT	TG	AAA	CAAG	TCA	CTG	CTC	CA	AA	ACCT	CTG	AC	T	1755
2y	2765	AGCAAG	CGG	AA	CCT	CTGG	AAA	AG	CAAT	CTG	TG	GA	TAA	GGTCA	2824
2b	1756	AGCAAG	CGG	AA	CCT	CTGG	AAA	AG	CAAT	CTG	TG	GA	TAA	GGTCA	1815
2y	2825	TGAG	CAG	TT	CC	AGAA	AG	ACCC	CAAG	CT	CAAG	CG	CAG	GA	2884
2b	1816	TGAG	CAG	TT	CC	AGAA	AG	ACCC	CAAG	CT	CAAG	CG	CAG	GA	1875
2y	2885	TAAAT	TCG	CT	CT	TTT	TG	TAT	TAT	TG	TAT	TACT	AAAT	TGG	2944
2b	1876	TAAAT	TCG	CT	CT	TTT	TG	TAT	TAT	TG	TAT	TACT	AAAT	TGG	1935
2y	2945	TTCAAT	AT	CC	CAG	TAA	ACCC	AT	TAT	TAT	TAT	TAT	TAT	TAT	3004

Qy	3005	TGTTCTAATGGTAAAGTACTGCGCTTTTGCACAGGAGCGCTGTTTCTTAAAGAAACCCATCGTGT	3064
Db	1996	TGTTCTAATGGTAAAGTACTGCGCTTTTGCACAGGAGCGCTGTTTCTTAAAGAAACCCATCGTGT	2055
Qy	3065	GAATATAGAGACTTTTCTACTGATCATCATAACTCTGTATCTGAGCAGTGTATACCAACCAC	3124
Db	2056	GAATATAGAGACTTTTCTACTGATCATCATAACTCTGTATCTGAGCAGTGTATACCAACCAC	2115
Qy	3125	ATCTGAAGTCAACAGAGAATCCAAAGTTTAAATTTGCTCGCGAATGTGTGCAGTATCTAG	3184
Db	2116	ATCTGAAGTCAACAGAGAATCCAAAGTTTAAATTTGCTCGCGAATGTGTGCAGTATCTAG	2175
Qy	3185	AAAAATGAACCGTAGTTTGTGTTTTTTTAAATACAGAAGTCATGTTGTTTCTTCTGCACTTTA	3244
Db	2176	AAAAATGAACCGTAGTTTGTGTTTTTTTAAATACAGAAGTCATGTTGTTTCTTCTGCACTTTA	2235
Qy	3245	TAATAAGCATGGAAGAAATTATCTTAGTAGGCAATTGTAAACACTTTTTCGAAAGTAAACC	3304
Db	2236	TAATAAGCATGGAAGAAATTATCTTAGTAGGCAATTGTAAACACTTTTTCGAAAGTAAACC	2295
Qy	3305	ATTTTCAGATTGAAATACTGCGATAATGTTGCTTTTAAAAAAGAAAAA	3354
Db	2296	ATTTTCAGATTGAAATACTGCAATAATGGTTGCTTTTAAAAAAGAAAAA	2345

RESULT 11

HUMLCAl

LOCUS

2198 bp mRNA linear PRI 18-DEC-2002

DEFINITION

Homo sapiens mRNA for l-caldesmon 1, complete cds.

ACCESSION

D90452

VERSION

D90452.1 GI:219895

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

Hayashi,K., Yano,H., Hashida,T., Takeuchi,R., Takeda,O., Asada,K., Takahashi,E., Kato,I. and Sobue,K.

TITLE

Genomic structure of the human caldesmon gene

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 89 (24), 12122-12126 (1992)

MEDLINE

93101679

PUBMED

1465449

REFERENCE

2 (bases 1 to 2198)

Hayashi,K. and Sobue,K.

AUTHORS

Direct Submission

TITLE

Submitted (26-JUL-1991) Kenji Sobue, Osaka University, Medical School, Biomedical Research Center, Department of Neurochemistry

JOURNAL

And Neuropharmacology; Kita-ku, Osaka 530, Japan

(Tel.:81-6-6445-6879, Fax:81-6-6445-6879)

ORIGIN

RESULT 12	
HUMCALD	linear
LOCUS	2975 bp mRNA
DEFINITION	Human caldesmon mRNA, complete cds.
ACCESSION	M64110
VERSION	M64110.1 GI:179829
KEYWORDS	F-actin binding protein; caldesmon; calmodulin; myosin; tropomyosin.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2975)
AUTHORS	Novy,R.E., Lin,J.L. and Lin,J.J.
TITLE	Characterization of cDNA clones encoding a human fibroblast caldesmon isoform and analysis of caldesmon expression in normal and transformed cells
JOURNAL	J. Biol. Chem. 266 (25), 16917-16924 (1991)
MEDLINE	91358497
PUBMED	1985618
COMMENT	Original source text: Homo sapiens (library: lambda gt11) fetus lung cDNA to mRNA.
FEATURES	Location/Qualifiers
source	1. .2975
	/organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="Wi-38"
/cell_type="fibroblast"
/tissue_type="lung"
/dev_stage="fetus"
/tissue_lib="lambda gt11"
112..1728
/codon_start=1
/product="caldesmon"
/protein_id="Aa35636.1"
/db_xref="GI:179830"
/translation="MDFFRRRLRQKREEMLEBRIAYQVNDDEEBAERRRR
ARQRLRQEEESLGQVTDQVNAQNSVDFDEAKTTNTQVGGDDDEAFLRLAR
EERQKRLQEALEKDFDPTTIDASLPSRQMDNTAENETTEKESESEQERY
EETETTVTKSQKNDWDAENKEDKEKEEKEPKRGSIGENQIKDKIKDKE
PKREVSKFMDRRKGFTEVSKNGCFWTHLKHTEFTSRPGGRASVTKGAEGAPQME
AKRLLELRREKGESEEPKQKQEALEELKKRERERKVLSEBEEQRKOE
EADRLREERERKLEETERRRAAEAKRQKMPEDGLSDDKPKPKFTPKGSLKIE
BRAFLNKSVQSKGVKSTHQAAIVSKIDSLRLEYTSIAIRGTSKAKTKPAASLPVP
AEGVRLKSMWKGWVFSPTAGTPEKTKAGLKVGVSSRLINEWLTPTDGNKSPAPK
PSDLRPGDVSSKRNLEKQSVKVTPTKV"

```

ORIGIN

```

Query Match      41.1%; Score 1522; DB 9; Length 2975;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1555; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1752 AGGAGAGAGAGAGGAACTAAAGTCGACGCTAAAGAGAGAAAGCTCCAGAGAGCAGC 1811
DB 653 AAGAAAACAAGAAAGAGACAGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 712
QY 1812 CTACCTTCAAAAAGAGAGATCAAGATGAAAGATTAAGAGAGAGAGAGAGAGAGAGAG 1871
DB 713 GGAGCATTTGGAGAGAAATCAGATCAAGATGAAAGATTAAGAGAGAGAGAGAGAGAG 772
QY 1872 AAGAGATTAAAGAGCTTCATGGATCGAAAGAGAGAGATTAAGAGAGAGAGAGAGAG 1931
DB 773 AAGAGATTAAAGAGCTTCATGGATCGAAAGAGAGAGATTAAGAGAGAGAGAGAGAGAG 832
QY 1932 GAGATTCTAGCCCAACAACTTAAACATACCTAGAGATCTTTTTCAGCGCCCTGGAGGGA 1991
DB 833 GAGATTCTAGCCCAACAACTTAAACATACCTAGAGATCTTTTTCAGCGCCCTGGAGGGA 892
QY 1992 GGGCCAGCGTGACACCAAGAGAGGCTGAGGCGCGCCCGAGAGAGAGAGAGAGAGAG 2051
DB 893 GGGCCAGCGTGACACCAAGAGAGGCTGAGGCGCGCCCGAGAGAGAGAGAGAGAGAG 952
QY 2052 TGGAGAGCTTCGTCCTCGCGGGAGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2111
DB 953 TGGAGAGCTTCGTCCTCGCGGGAGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1012
QY 2112 AGAAGCAGCAGAGAGCGGCTTTGGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2171
DB 1013 AGAAGCAGCAGAGAGCGGCTTTGGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1072
QY 2172 GGAAGCTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2231
DB 1073 GGAAGCTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1132
QY 2232 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2291
DB 1133 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1192
QY 2292 CTGAGAAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2351
DB 1193 CTGAGAAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1252
QY 2352 GTTTCACCTCTAAAGGTTCACTCTCAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2411
DB 1253 GTTTCACCTCTAAAGGTTCACTCTCAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1312
QY 2412 CTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2471

```

```

DB 1313 CTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1372
QY 2472 ACAGCAGAGCTGGAGCAGATATACCAAGTCGAATTTGAGGGGAGAGAGAGAGAGAGAGAG 2531
DB 1373 ACAGCAGAGCTGGAGCAGATATACCAAGTCGAATTTGAGGGGAGAGAGAGAGAGAGAGAG 1432
QY 2532 AGCCGGCAGAGCTCGGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2591
DB 1433 AGCCGGCAGAGCTCGGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1492
QY 2592 GGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2651
DB 1493 GGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1552
QY 2652 CTGGCTTGAAGGTAGGGGTTTCTAGCGCGCATCAATGAATGCTTAACTAAACCCAGATG 2711
DB 1553 CTGGCTTGAAGGTAGGGGTTTCTAGCGCGCATCAATGAATGCTTAACTAAACCCAGATG 1612
QY 2712 GAAACCAAGTCACCTGCTCCCAACCTTCTGACTTGGAGAGAGAGAGAGAGAGAGAGAGAG 2771
DB 1613 GAAACCAAGTCACCTGCTCCCAACCTTCTGACTTGGAGAGAGAGAGAGAGAGAGAGAGAG 1672
QY 2772 GGAACCTCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2831
DB 1673 GGAACCTCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1732
QY 2832 GTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2891
DB 1733 GTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1792
QY 2892 CTCTGTTTTGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2951
DB 1793 CTCTGTTTTGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1852
QY 2952 TCCAGTAAACCCATGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3011
DB 1853 TCCAGTAAACCCATGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1912
QY 3012 GGTAAAGTACTGCTTTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3071
DB 1913 GGTAAAGTACTGCTTTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1972
QY 3072 AGACTTTTCTACTGATCATCAATACTCTGATCTGAGCAGTGATACCAACCACTCTGAA 3131
DB 1973 AGACTTTTCTACTGATCATCAATACTCTGATCTGAGCAGTGATACCAACCACTCTGAA 2032
QY 3132 GTCAACAGAGATCCCAAGTTTAAATTTGCTGCGGAAATGTGCGAGATCTAGAGAGAGAG 3191
DB 2033 GTCAACAGAGATCCCAAGTTTAAATTTGCTGCGGAAATGTGCGAGATCTAGAGAGAGAG 2092
QY 3192 AACCGTAGTTTTGTTTTTAAATACAGAGATCATGTTGTTTCTGCACTTTTATATAA 3251
DB 2093 AACCGTAGTTTTGTTTTTAAATACAGAGATCATGTTGTTTCTGCACTTTTATATAA 2152
QY 3252 GCATGGAAGAGATTTATCTTAGTAGGCAATTTGTAACACTTTTGAAGTAAACCACTTCAG 3311
DB 2153 GCATGGAAGAGATTTATCTTAGTAGGCAATTTGTAACACTTTTGAAGTAAACCACTTCAG 2212
QY 3312 ATTTGAAATCTCCGATATGTTGCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3361
DB 2213 ATTTGAAATCTCCGATATGTTGCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2262

```

```

RESULT 13
LOCUS AR001262
DEFINITION Sequence 13 from patent US 5739008.
ACCESSION AR001262
VERSION AR001262.1
KEYWORDS GI:3963329
SOURCE Unknown.
ORGANISM Unknown.

```

1375	GGGAGAAAGGGAATGTGTTTTTCATCCCCCACTGCACGGCACACCAATTAAGGAAATCG	1434
2652	CTGGCTTGAAGTAGGGGTTTCTAGCCGATCAATGAATGGCTAACTAAACCCCGATG	2711
1435	CTGGCTTGAAGTAGGGGTTTCTAGCCGATCAATGAATGGCTAACTAAACCCCGATG	1494
2712	GAACAAGTCACTGCTCCCAACCTTCTGACTTGAGACGAGGAGCGTATCCAGCAAGC	2771
1495	GAACAAGTCACTGCTCCCAACCTTCTGACTTGAGACGAGGAGCGTATCCAGCAAGC	1554
2772	GGAACTCTGGGAAAGCAATCTGTGGATAAGTCACTTCCCCCACTTAAGGTTTGAGACA	2831
1555	GGAACTCTGGGAAAGCAATCTGTGGATAAGTCACTTCCCCCACTTAAGGTTTGAGACA	1614
2832	GTTCAGAAAGAACCCAAAGCTCAAGACGAGGAGCGTCACTGTGTAGGGGCTAATTCG	2891
1615	GTTCAGAAAGAACCCAAAGCTCAAGACGAGGAGCGTCACTGTGTAGGGGCTAATTCG	1674
2892	CTCTGTTTGTATTTATGTTGTTTAACTAAATTTGGGTTCAATATCTTTTATTTTCAATA	2951
1675	CTC--TTTGTATTTATGTTGTTTAACTAAATTTGGGTTCAATATCTTTTATTTTCAATA	1732
2952	TCCAGTAAACCCATGTATTTATCACTATATTTAATAATCAAGTCTAGAGATGTTTCAT	3011
1733	TCCAGTAAACCCATGTATTTATCACTATATTTAATAATCAAG--CTAGAGATGTTTCAT	1791
3012	GGTAAAGTACTGCCCTTTCACAGGAGCGCTTTCTAAAGAAACCCCATGCTGTGAAATAG	3071
1792	GGTAAAGTACTGCCCTTTCACAGGAGCGCTTTCTAAAGAAACCCCATGCTGTGAAATAG	1851
3072	AGACTTTTCTACTGATCATCACTCTGATCTGAGCAGTGATACCAACCATCTGAA	3131
1852	AGACTTTTCTACTGATCATCACTCTGATCTGAGCAGTGATACCAACCATCTGAA	1911
3132	GTCAACAGAGATGCCAAGTTTAAATTTGCCGCGGAATGTGTCAGTATCTAGAAAAATG	3191
1912	GTCAACAGAGATGCCAAGTTTAAATTTGCCGCGGAATGTGTCAGTATCTAGAAAAATG	1971
3192	AACCGTAGTTTTGTTTTTAAATCAGAAAGTCATGTTGTTCTGACATTTATAATAAA	3251
1972	AACCGTAG--TTTTGTTTTTAAATCAGAAAGTCATGTTGTTCTGACATTTATAATAAA	2030
3252	GCATGGAAGAAATATCTTAGTAGGCAATTTGAACACTTTTTGAAAGTAACCCATTTGAG	3311
2031	GCATGGAAGAAATATCTTAGTAGGCAATTTGAACACTTTTTGAAAGTAACCCATTTGAG	2090
3312	ATTGGAATCTCGATATGTTGTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA	3357
2091	ATTGGAATCTCGAATAATGTTGTTTAAATTTAAATTTAAATTTAAATTTAAATTTAA	2136
RESULT 14		
123420		
LOCUS	2137 bp DNA linear PAT 07-OCT-1999	
DEFINITION	Sequence 13 from patent US 5532337.	
ACCESSION	123420	
VERSION	123420.1	
KEYWORDS	GI:1603290	
SOURCE	Unknown.	
ORGANISM	Unclassified.	
REFERENCE	1 (bases 1 to 2137)	
AUTHORS	Hayashi,K., Hashida,T., Asada,K., Kotani,H., Kato,I. and Sobue,K.	
TITLE	Proteins comprising calmodulin- and actin-binding human caldesmon	
JOURNAL	polypeptide fragments	
FEATURES	Patent: US 5532337-A 13 02-JUL-1996;	
source	Location/Qualifiers	
	1..2137	
	/organism="unknown"	
ORIGIN	/mol_type="unassigned DNA"	

Db	1555	GGAACTCTGGGAAAGCAATCTGTGATAAGTCTACTCTCCCACTAAGGTTTGAGACA	1614
Qy	2832	GTTCGAGAAAGAACCAAGCTCAAGACGACGACGAGCTCAGTTGTAGAGGGCTAATTCG	2891
Db	1615	GTTCGAGAAAGAACCAAGCTCAAGACGACGACGAGCTCAGTTGTAGAGGGCTAATTCG	1674
Qy	2892	CTCTGTTTTTGTAATTTATGTTGATTACTAAATTTGGGTTTCATTATCTTTTATTTTCAATA	2851
Db	1675	CTC--TTTGTATTTATGTTGATTACTAAATTTGGGTTTCATTATCTTTTATTTTCAATA	1732
Qy	2952	TCCAGTAAACCCATGTATATATCACTATATTTAATAATCACAGTCTAGAGATGTTTCAT	3011
Db	1733	TCCAGTAAACCCATGTATATATCACTATATTTAATAATCACAG-CTAGAGATGTTTCAT	1791
Qy	3012	GGTAAAGTACTGCCTTTGCACAGGAGCCCTGTTTCTAAAGAAACCCATGCTGTGAATAG	3071
Db	1792	GGTAAAGTACTGCCTTTGCACAGGAGCCCTGTTTCTAAAGAAACCCATGCTGTGAATAG	1851
Qy	3072	AGACTTTCTACTGATCATCATTAACCTGTGTATCTGACAGTGATACCAACACACTCTGAA	3131
Db	1852	AGACTTTCTACTGATCATCATTAACCTGTGTATCTGACAGTGATACCAACACACTCTGAA	1911
Qy	3132	GTCAACAGAAAGATCCAAAGTTTAAATTTGCCCTGGGAAATGTTGTCAGTATCTAGAAAAATG	3191
Db	1912	GTCAACAGAAAGATCCAAAGTTTAAATTTGCCCTGGGAAATGTTGTCAGTATCTAGAAAAATG	1971
Qy	3192	AACGGTAGTTTTGTTTTTAAATACAGAGTCACTGCTTCTGTCACATTTATAATAAA	3251
Db	1972	AACGGTAG-TTTTGTGTTTTTAAATACAGAGTCACTGCTTCTGTCACATTTATAATAAA	2030
Qy	3252	GCATGGAGAAATTTATCTTAGTAGGCAATTTGAACACTTTTGAAGTAACCCATTTTCAG	3311
Db	2031	GCATGGAGAAATTTATCTTAGTAGGCAATTTGAACACTTTTGAAGTAACCCATTTTCAG	2090
Qy	3312	ATTTGAATACTCGATAATGGTTGCTTTTAAABAAAAGAAAAGA	3357
Db	2091	ATTTGAATACTCGATAATGGTTGCTTTTAAABAAAAGAAAAGA	2136
RESULT 15			
HUMLCA2			
LOCUS	HUMLCA2	2120 bp	mRNA linear PRI 18-DEC-2002
DEFINITION	Homo sapiens mRNA for l-caldesmon II, complete cds.		
ACCESSION	D90453		
VERSION	D90453.1	GI:219897	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Hayashi, K., Yano, H., Hashida, T., Takeuchi, R., Takeda, O., Asada, K., Takahashi, E., Kato, I. and Sobue, K.		
TITLE	Genomic structure of the human caldesmon gene		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 89 (24), 12122-12126 (1992)		
MEDLINE	93101679		
REFERENCE	2		
PUBMED	1465449		
	(bases 1 to 2120)		
AUTHORS	Hayaishi, K. and Sobue, K.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-JUL-1991) Kenji Sobue, Osaka University, Medical School, Biomedical Research Center, Department of Neurochemistry and Neuropharmacology, Kita-Ku, Osaka 530, Japan		
	(Tel:81-6-6445-6879, Fax:81-6-6445-6879)		
FEATURES	Location/Qualifiers		
source	1..2120		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/cell_line="HeLa S3"		
	12..1610		
	/codon_start=1		
CDS			

```
/product="l-caldesmon II"
/protein_id="BAA14419.1"
/db_xref="GI:219898"
/translations="MLGSGSHGRSLAALSQIAYQNRDDDEEBAARRRRARRQERL
ROKQBEESLGQVTDVEVNAQNSVPDEEAKTTTNTQVEGDDDEAAFLERLARRBERRO
KSLQEALESQKEDFDPTIDALSLSIPSRMONDPAENETTEKEESQSERVEIEETE
TVTKSVQKNDWRDAENKEDSKKEBEEKPKGSGIGENQIKDEKIKKDEKPEKEEVX
SMDRKKGFTEVSKQNGFEWTHKDKHTENTFSRPGGRASVDTKBAEGAPQVEAGKRL
ELURRGEETSEFEKLKQOEAALEELKKREBRKRVLEEEERKQKEEEDARFL
REEEERRLKEEIERRAEAERKQMPEDGLSDKKPFKCFKTFKSGSLKIEERABFL
NKSQVSSGVKSTHQAIIVSKIDSLRLEQVTSAREGTSKAPTKPAASPLPVPABGVN
IKMSWKGNVFSGPTAAAGTGNKETAAGLKVGVSSRINELWLTPTDNKSPAPKPSDLRP
GDVSSXENLWEXQSDVKVTSPTK"
```

ORIGIN

Query Match	39.7%;	Score	1469.4;	DB	9;	Length	2120;
Best Local Similarity	96.5%;	Pred. No.	6.6e-304;				
Matches	1534;	Conservative	0;	Mismatches	51;	Indels	4;
				Gaps	3;		
QY	1752	AGGGAGAAGAGAGGGAACCTAAAGTGAAGCTAAAGAGAGAAAAGCTCCAAAGAGACAAGC	1811				
DB	535	AGAAAACAAGAAAGAAAGACACAGAAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	594				
QY	1812	CTACCTTCAAAAAGAGAGATCAAGATGAAAGATTAAGATTAAGAGAGAGAGAGAGAGAGAGAG	1871				
DB	595	GGAGCATTCGAGAAAATCAGATCAAGATCAAGATTAAGATTAAGAGAGAGAGAGAGAGAGAG	654				
QY	1872	AGAGATTAAGAGCTTCATCGATCGAAGAGAGGATTTACAGAGCTTAAGTCGCAGAGATG	1931				
DB	655	AGAGATTAAGAGCTTCATCGATCGAAGAGAGGATTTACAGAGCTTAAGTCGCAGAGATG	714				
QY	1932	GAGAATTCATGACCCACAAACTTAAACATATCTAGAGATTAAGTCAGCGCCCTCGAGGGA	1991				
DB	715	GAGAATTCATGACCCACAAACTTAAACATATCTAGAGATTAAGTCAGCGCCCTCGAGGGA	774				
QY	1992	GGCCCGAGTGGACCAAGAGGCTGAGGCGCCCGCCAGGTGAAGCGGCAAGAGGC	2051				
DB	775	GGCCCGAGTGGACCAAGAGGCTGAGGCGCCCGCCAGGTGAAGCGGCAAGAGGC	834				
QY	2052	TGAGGAGCTTCGTCGTGTCGCGGGGAGACCGAGAGCGAAGAGTTCGAGAGCTCAAAAC	2111				
DB	835	TGAGGAGCTTCGTCGTGTCGCGGGGAGACCGAGAGCGAAGAGTTCGAGAGCTCAAAAC	894				
QY	2112	AGAAGCAGAGGAGCGGCTTGAGCTGAGGAGCTCAAGAGAAAGAGGAGGAGAGAA	2171				
DB	895	AGAAGCAGAGGAGCGGCTTGAGCTGAGGAGCTCAAGAGAAAGAGGAGGAGAGAA	954				
QY	2172	GGAAGTCTCTGGAGGAG	2231				
DB	955	GGAAGTCTCTGGAGGAG	1014				
QY	2232	GAGAGGAG	2291				
DB	1015	GAGAGGAG	1074				
QY	2292	CTGAGAAACGCGCAG	2351				
DB	1075	CTGAGAAACGCGCAG	1134				
QY	2352	GTTTCACTCTCTAAAGGTTTCATCTCTCAAGATAGAGAGAGAGAGAGAGAGAGAG	2411				
DB	1135	GTTTCACTCTCTAAAGGTTTCATCTCTCAAGATAGAGAGAGAGAGAGAGAGAGAG	1194				
QY	2412	CTGTGCAAGAAAGCAGTGGTGTCAAAATCGACCCATCAAGCAGCAATAGTCTCCAGATTG	2471				
DB	1195	CTGTGCAAGAAAGCAGTGGTGTCAAAATCGACCCATCAAGCAGCAATAGTCTCCAGATTG	1254				
QY	2472	ACAGCAGACTGGAGCAGTATACAGTGCATTTGAGGGAGACAAAGCGCAAAACCTACAA	2531				
DB	1255	ACAGCAGACTGGAGCAGTATACAGTGCATTTGAGGGAGACAAAGCGCAAAACCTACAA	1314				
QY	2532	AGCCGGCAGCCTCGATCTCTCTGTTCTGCTGAAGGTTGACGCAACATCAAGAGTATGT	2591				

DB	1315	AGCCGCGAGCCTCGGATCTTCCTGTTCTGCTGAAGGTGAGCAACATCAAGAGTATGT	1374				
QY	2592	GGGAGAAAGGGAATGTGTTTTCATCCCCCACTGCGAGGAGCACACCAATTAAGAAACTG	2651				
DB	1375	GGGAGAAAGGGAATGTGTTTTCATCCCCCACTGCGAGGAGCACACCAATTAAGAAACTG	1434				
QY	2652	CTGGCTTGAAGGTAGGGGTTTCTAGCGGCATCAATGAATGCTTAATAAAACCCACAGATG	2711				
DB	1435	CTGGCTTGAAGGTAGGGGTTTCTAGCGGCATCAATGAATGCTTAATAAAACCCACAGATG	1494				
QY	2712	GAAACAAGTCACTGCTGCCAAACCTTCTGACTTGAAGCCAGGAGACGTATCCAGCAAGC	2771				
DB	1495	GAAACAAGTCACTGCTGCCAAACCTTCTGACTTGAAGCCAGGAGACGTATCCAGCAAGC	1554				
QY	2772	GGAACTCTGGGAAAGCAATCTGTGATAGGTCACCTTCCCCCACTAAGGTTTTCAGACA	2831				
DB	1555	GGAACTCTGGGAAAGCAATCTGTGATAGGTCACCTTCCCCCACTAAGGTTTTCAGACA	1614				
QY	2832	GTTCCGAAAGAACCCCAAGCTCAAGCGCAGGAGCTCAGTTGTAGAGGCGCTAATTCG	2891				
DB	1615	GTTCCGAAAGAACCCCAAGCTCAAGCGCAGGAGCTCAGTTGTAGAGGCGCTAATTCG	1674				
QY	2892	CTCTGTTTGTATTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATG	2951				
DB	1675	CTC--TTTGTATTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATG	1732				
QY	2952	TCCAGGTAAACCCATGTATTTATCTACTATATTTAATAATCACAGTCTAGAGATGTTCA	3011				
DB	1733	TCCAGGTAAACCCATGTATTTATCTACTATATTTAATAATCACAG-CTAGAGATGTTCA	1791				
QY	3012	GGTAAAGTACTGCTTTTGACAGGAGCGCTTTTCTAAGAAACCCCATGCTGTGAAATAG	3071				
DB	1792	GGTAAAGTACTGCTTTTGACAGGAGCGCTTTTCTAAGAAACCCCATGCTGTGAAATAG	1851				
QY	3072	AGACTTTTCTACTGATCATCATAACTCTGTATCTGAGCAGTGATACCAACCATCTGAA	3131				
DB	1852	AGACTTTTCTACTGATCATCATAACTCTGTATCTGAGCAGTGATACCAACCATCTGAA	1911				
QY	3132	GTCAACAGAGATCCAGTTTAAATTTGCTGCGGAAATGCTGCGAGTATCTAGAAAATG	3191				
DB	1912	GTCAACAGAGATCCAGTTTAAATTTGCTGCGGAAATGCTGCGAGTATCTAGAAAATG	1971				
QY	3192	AACCGTAGTTTTTGTGTTTTTAAATACAGAAAGTCAATGTTGTTTCTGCACTTTATAATAA	3251				
DB	1972	AACCGTAG-TTTTGTGTTTTTAAATACAGAAAGTCAATGTTTCTGCACTTTATAATAA	2030				
QY	3252	GCATGGAAGAAATTTATCTTAGTAGGCAATTTGACACTTTTGAAGTAAACCATTTTCAG	3311				
DB	2031	GCATGGAAGAAATTTATCTTAGTAGGCAATTTGACACTTTTGAAGTAAACCATTTTCAG	2090				
QY	3312	ATTTGAAATACTGCGATAATGGTTGCTTT	3340				
DB	2091	ATTTGAAATACTGCGATAATGGTTGCTTT	2119				

Search completed: March 11, 2004, 03:36:40
Job time : 9450.45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 08:01:57 ; Search time 900.994 Seconds
(without alignments)
17454.974 Million cell updates/sec

Title: US-10-084-817-30

Perfect score: 3702

Sequence: 1 gcctagggaatgacaggcat.....aaggggcggtaccgagaac 3702

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124039041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001as.*
- 5: Geneseq2001bs.*
- 6: Geneseq2002s.*
- 7: Geneseq2003as.*
- 8: Geneseq2003bs.*
- 9: Geneseq2003cs.*
- 10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3115.4	84.2	3610	6	ABL62384 Colon ade
2	3116.4	84.2	3610	6	ABL68604 Kidney ca
3	3116.4	84.2	3610	6	ABN95909 Gene #240
4	1793.6	48.4	4547	7	ABX77534 Different
5	1567	42.3	2215	2	AAQ28988 Sequence
6	1455.8	39.3	3342	4	AAH23102 Osteoartr
7	1452.8	39.3	2137	2	AAQ28987 Sequence
8	1086.2	29.3	1677	2	AAQ28986 Sequence
9	999.2	27.0	1599	2	AAQ28985 Sequence
10	932.6	25.2	539	2	AAQ28984 Sequence
11	691	18.7	705	3	AAC00905 Human sec
12	689.2	18.6	756	2	AAH39869 Gastric c
13	597.6	16.1	706	2	AAH39870 Gastric c
14	584.2	15.8	2511	3	AAH39870 Human can
15	584.2	15.8	2511	4	AAH33117 Human col
16	546.8	14.7	2209	7	ABX77533 Different
17	544.8	14.7	548	3	AAC01571 Human sec
18	488	13.1	1750	2	AAQ23748 1-Caldesm
19	479.8	13.0	1014	2	AAQ12340 Caldesmon
20	447.4	12.1	449	6	ABV95786 Human pan
21	436.6	11.8	503	8	ACH32073 Human end
22	428.6	11.6	436	8	ACH15816 Human adu
23	428.6	11.6	606	3	AAAL6290 Human col

C

ALIGNMENTS

RESULT 1

ABL62384

ID ABL62384 standard; DNA; 3610 BP.

XX ABL62384;

AC ABL62384;

XX 15-MAY-2002 (first entry)

DT 15-MAY-2002 (first entry)

XX 15-MAY-2002 (first entry)

DE Colon adenocarcinoma related gene sequence SEQ ID NO:721.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

KW stomach; lung; prostate; pancreas; carcinoma; antitumour; carcinos;

KW cytosolic; Gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;

XX Gene; ds.

XX Homo sapiens.

OS Homo sapiens.

XX WO200194629-A2.

PN WO200194629-A2.

XX 13-DEC-2001.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US010838.

XX 05-JUN-2000; 2000US-0209473P.

XX 05-JUN-2000; 2000US-0209531P.

XX 18-SEP-2000; 2000US-0233133P.

XX 18-SEP-2000; 2000US-0233617P.

XX 20-SEP-2000; 2000US-0234009P.

XX 20-SEP-2000; 2000US-0234034P.

XX 20-SEP-2000; 2000US-0234052P.

XX 22-SEP-2000; 2000US-0234509P.

XX 22-SEP-2000; 2000US-0234567P.

XX 25-SEP-2000; 2000US-0234923P.

XX 25-SEP-2000; 2000US-0235077P.

XX 25-SEP-2000; 2000US-0235082P.

XX 25-SEP-2000; 2000US-0235134P.

XX 25-SEP-2000; 2000US-0235280P.

XX 26-SEP-2000; 2000US-0235637P.

XX 26-SEP-2000; 2000US-0235638P.

XX 27-SEP-2000; 2000US-0235711P.

XX 27-SEP-2000; 2000US-0235720P.

XX 27-SEP-2000; 2000US-0235840P.

XX 27-SEP-2000; 2000US-0235863P.

XX 28-SEP-2000; 2000US-0236028P.

XX 28-SEP-2000; 2000US-0236032P.

ABV55814 Human pro
ABN74587 Bovine em
AAI90712 Human pol
AAQ28983 Sequence
AAQ28982 Sequence
AAQ28981 Sequence
AAI25076 Human bre
ABL75838 Corn tass
AAQ28998 Sequence
AAQ28996 Sequence
AAC78295 Human can
ABQ55690 Human ova
ABX38333 Bovine ES
ABX38333 Sequence
ABL05274 Drosophil
AAL0424 Human sec
AAT05868 Chicken 1
AAD21685 Mutational
AAD21684 Human ret
AAH70099 Plasmidli
ABN65775 Human can
ABN65775 Human can
ABL05275 Drosophil

Db ||||| 1321 AAGGCGCAGCAGAGGAGGCGCAAGGCGCCAGGCGCAGAGGAGGAGGCGCTAAGGTAGA 1380
2Y AGACGAGAAAAGCTTAAACAGCAGCTAGAGAGAGAAAAACATGCGCATCCAAAGAGACAAAGAT 1656
Db CCGCATCAATGAATGGCTAACTAAACCCAGATGGAACCAAGTCACCTGCTCCCAAAC 2520
2Y TTCTGACTTGAGACCCAGAGAGCGTATCCAGCAAGCGGAACCTCTGGGAAAAGCAATCTGT 2796
Db TTCTGACTTGAGACCCAGAGAGCGTATCCAGCAAGCGGAACCTCTGGGAAAAGCAATCTGT 2580
2Y GGATAAGGTCACTTCCGCCACCTAAAGTTTGGACAGCTTCCAGAAAAGCAACCAAGCTCAAG 2856
Db GGATAAGGTCACTTCCGCCACCTAAAGTTTGGACAGCTTCCAGAAAAGCAACCAAGCTCAAG 2640
2Y ACGCAGGACGAGCTCAGTTTGTAGAGGCTAAATTCGCTCTGTTTGTATTTATGTTGATTT 2916
Db ACGCAGGACGAGCTCAGTTTGTAGAGGCTAAATTCGCTCTGTTTGTATTTATGTTGATTT 2700
2Y ACTAAATGGGTTCAATATCTTTTATTTTCAATATCCCAAGTAAACCCATGTATATTC 2976
Db ACTAAATGGGTTCAATATCTTTTATTTTCAATATCCCAAGTAAACCCATGTATATTC 2760
2Y ACTATATTTAATAATCACAGTCTAGAGATGTTTCATGTTAAAGTACTGCTTTCACACAGG 3036
Db ACTATATTTAATAATCACAGTCTAGAGATGTTTCATGTTAAAGTACTGCTTTCACACAGG 2820
2Y AGCCTGTTTCTAAAGAAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCAATAC 3096
Db ATCCTGTTTCTAAAGAAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCAATAC 2880
2Y TCTGATCTGAGCAGTGATACCAACCACTCTGAAAGTCAACAGAGATCCAAGTTTAAAA 3156
Db TCTGATCTGAGCAGTGATACCAACCACTCTGAAAGTCAACAGAGATCCAAGTTTAAAA 2940
2Y TTGCTCGGGAATGTGTCAGTATCTAGAAAAGTAAACCGTAGTTTGTGTTTAAAAAT 3216
Db TTG-CTGCGGAATGTGTCAGTATCTAGAAAAGTAAACCGTAGTTTGTGTTTAAAAAT 2999
2Y ACAGAGTCATGTTGTTTCTGCACTTTATAATAAGCATGGAAGAAATTTATCTTAGTAGG 3276
Db ACAGAGTCATGTTGTTTCTGCACTTTATAATAAGCATGGAAGAAATTTATCTTAGTAGG 3059
2Y CAATGTAAACACTTTTGAAGTAACCCATTTAGAAATTTGAAATTTGAAATTTGAAATTTG 3336
Db CAATGTAAACACTTTTGAAGTAACCCATTTAGAAATTTGAAATTTGAAATTTGAAATTTG 3119
2Y CTTTAAAAAAGAAAAGATGCGCGCAGTATTTCCTTT 3378
Db TCTTTAAAAAAGAAAAGATGCTTAAAGTATTACTTTT 3161

RESULT 2
ABL68604
ID ABL68604 standard; DNA; 3610 BP.
XX ABL68604;
AC ABL68604;
DT 15-MAY-2002 (first entry)
XX
DE Kidney cancer related gene sequence SEQ ID NO:6941.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.

Db ||||| 1321 AAGGCGCAGCAGAGGAGGCGCAAGGCGCCAGGCGCAGAGGAGGAGGCGCTAAGGTAGA 1380
2Y AGACGAGAAAAGCTTAAACAGCAGCTAGAGAGAGAAAAACATGCGCATCCAAAGAGACAAAGAT 1656
Db CCGCATCAATGAATGGCTAACTAAACCCAGATGGAACCAAGTCACCTGCTCCCAAAC 2520
2Y TTCTGACTTGAGACCCAGAGAGCGTATCCAGCAAGCGGAACCTCTGGGAAAAGCAATCTGT 2796
Db TTCTGACTTGAGACCCAGAGAGCGTATCCAGCAAGCGGAACCTCTGGGAAAAGCAATCTGT 2580
2Y GGATAAGGTCACTTCCGCCACCTAAAGTTTGGACAGCTTCCAGAAAAGCAACCAAGCTCAAG 2856
Db GGATAAGGTCACTTCCGCCACCTAAAGTTTGGACAGCTTCCAGAAAAGCAACCAAGCTCAAG 2640
2Y ACGCAGGACGAGCTCAGTTTGTAGAGGCTAAATTCGCTCTGTTTGTATTTATGTTGATTT 2916
Db ACGCAGGACGAGCTCAGTTTGTAGAGGCTAAATTCGCTCTGTTTGTATTTATGTTGATTT 2700
2Y ACTAAATGGGTTCAATATCTTTTATTTTCAATATCCCAAGTAAACCCATGTATATTC 2976
Db ACTAAATGGGTTCAATATCTTTTATTTTCAATATCCCAAGTAAACCCATGTATATTC 2760
2Y ACTATATTTAATAATCACAGTCTAGAGATGTTTCATGTTAAAGTACTGCTTTCACACAGG 3036
Db ACTATATTTAATAATCACAGTCTAGAGATGTTTCATGTTAAAGTACTGCTTTCACACAGG 2820
2Y AGCCTGTTTCTAAAGAAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCAATAC 3096
Db ATCCTGTTTCTAAAGAAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCAATAC 2880
2Y TCTGATCTGAGCAGTGATACCAACCACTCTGAAAGTCAACAGAGATCCAAGTTTAAAA 3156
Db TCTGATCTGAGCAGTGATACCAACCACTCTGAAAGTCAACAGAGATCCAAGTTTAAAA 2940
2Y TTGCTCGGGAATGTGTCAGTATCTAGAAAAGTAAACCGTAGTTTGTGTTTAAAAAT 3216
Db TTG-CTGCGGAATGTGTCAGTATCTAGAAAAGTAAACCGTAGTTTGTGTTTAAAAAT 2999
2Y ACAGAGTCATGTTGTTTCTGCACTTTATAATAAGCATGGAAGAAATTTATCTTAGTAGG 3276
Db ACAGAGTCATGTTGTTTCTGCACTTTATAATAAGCATGGAAGAAATTTATCTTAGTAGG 3059
2Y CAATGTAAACACTTTTGAAGTAACCCATTTAGAAATTTGAAATTTGAAATTTGAAATTTG 3336
Db CAATGTAAACACTTTTGAAGTAACCCATTTAGAAATTTGAAATTTGAAATTTGAAATTTG 3119
2Y CTTTAAAAAAGAAAAGATGCGCGCAGTATTTCCTTT 3378
Db TCTTTAAAAAAGAAAAGATGCTTAAAGTATTACTTTT 3161

RESULT 2
ABL68604
ID ABL68604 standard; DNA; 3610 BP.
XX ABL68604;
AC ABL68604;
DT 15-MAY-2002 (first entry)
XX
DE Kidney cancer related gene sequence SEQ ID NO:6941.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.

CC	tumour
XX	Sequence 3610 BP; 1322 A; 622 C; 963 G; 703 T; 0 U; 0 Other;
XX	Query Match 84.2%; Score 3116.4; DB 6; Length 3610; Best Local Similarity 99.3%; Pred. No. 0; Matches 3140; Conservative 0; Mismatches 21; Indels 1; Gaps 1;
QY	217 CCTGCTGCCCGCCTGTCTCTCTGGCTGTGCTCTCTTAAAGAAATCAAGTCCTTTCC 276 DB 1 CCTGCTGCCCGCCTGTCTCTCTGGCTGTGCTCTCTTAAAGAAATCAAGTCCTTTCC 60
QY	277 TTTCGACTTAGTCTCTCGGAAGAAGTTTACAGTACAAGGTATCATTTGGACATTTCAA 336 DB 61 TTTCGACTTAGTCTCTCGGAAGAAGTTTACAGTACAAGGTATCATTTGGACATTTCAA 120
QY	337 GATCATCAAATCAAATTCACAGGGATTTGGTGCACCACCAAGAGGTTCAGACATCTGATT 396 DB 121 GATCATCAAATCAAATTCACAGGGATTTGGTGCACCACCAAGAGGTTCAGACATCTGATT 180
QY	397 GTTGACCTGTCAGACATCATCTGTCTCTCTTGAACCTGAATATCACACCATGATGATTT 456 DB 181 GTTGACCTGTCAGACATCATCTGTCTCTCTTGAACCTGAATATCACACCATGATGATTT 240
QY	457 TGAGCGTTCGACAGAACTTTAGAAGCAAAAGAGGAGGAGATGCGACTCGAAGCAAAAAG 516 DB 241 TGAGCGTTCGACAGAACTTTAGAAGCAAAAGAGGAGGAGATGCGACTCGAAGCAAAAAG 300
QY	517 AATCGCTACACAGAGAAATGACATGATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 576 DB 301 AATCGCTACACAGAGAAATGACATGATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY	577 AGCCCGACAGGAAACGGCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 636 DB 361 AGCCCGACAGGAAACGGCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY	637 CCAGGTGGAGGTGAATGCCAGAACAGTGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 696 DB 421 CCAGGTGGAGGTGAATGCCAGAACAGTGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY	697 AAACACTCAAGTGGAAAGGGAGTATGAGGCGCGCATTTCTTGAGAGCGCTGCTGCGCGTGA 756 DB 481 AAACACTCAAGTGGAAAGGGAGTATGAGGCGCGCATTTCTTGAGAGCGCTGCTGCGCGTGA 540
QY	757 GGAAGACGCGCAAAACGGCTTCAGGAGGCTTCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAG 816 DB 541 GGAAGACGCGCAAAACGGCTTCAGGAGGCTTCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY	817 AATAACAGATGCAAGTCTGTCTGCTCCCAAGCAGAGGAATCCAAATGACACAGCAGAAAA 876 DB 601 AATAACAGATGCAAGTCTGTCTGCTCCCAAGCAGAGGAATCCAAATGACACAGCAGAAAA 660
QY	877 TGAATCTCCGAGAGGAGAGAAAAGTAAGTCCGCAAGAAAGATACAGATAGAGGA 936 DB 661 TGAATCTCCGAGAGGAGAGAAAAGTAAGTCCGCAAGAAAGATACAGATAGAGGA 720
QY	937 AACAGAAACAGTCCCAAGTCTTACCAAGAAATGATTGGAGGATGCTGGAAGAAAAACA 996 DB 721 AACAGAAACAGTCCCAAGTCTTACCAAGAAATGATTGGAGGATGCTGGAAGAAAAACA 780
QY	997 GAAGAGACAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1056 DB 781 GAAGAGACAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
QY	1057 AGAAATCAGGTAGAGTGTGTGGAGAGAAAAACAATCTGAAGCCAGGAGGAGGAGGAGGAGT 1116 DB 841 AGAAATCAGGTAGAGTGTGTGGAGAGAAAAACAATCTGAAGCCAGGAGGAGGAGGAGT 900
QY	1117 GGTAAATCTATTAATAATGGGAGATCAGTTCAGAGAGGCTTAAACAAGAGGAGGAGAG 1176 DB 901 GGTAAATCTATTAATAATGGGAGATCAGTTCAGAGAGGCTTAAACAAGAGGAGGAGAG 960
QY	1177 GGAAACAGGTTCAAGATGAGATTTCCCATCATGAAGATGAAGAGAGGAGAGCAAGAGAAAG 1236

(AVAIL-) AVOLON PHARM.
Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,
Soppet DR, Weaver Z;
WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.

Claim 1; SEQ ID NO 6941; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing calls to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' cell carcinoma.

[illegible]

1620 TAGAGAGAGAAAAAATCATGCCATGCAAGAGAGCAAAAGATAAAGGGGAAAAAGGTAGAACAGA 1679
 1149 ----- 1148
 1680 AATAGAGGGAATGGTAAATGAAGAAAGAAAGCAAGAGAGATAAACTTCAGACAGCTG 1739
 1149 ----- 1148
 1740 TCCTAAAGAAACAGGGAGAGAGAGAGGGAACATAAAGTGCAGCTAAAGAGAGAAAGCTCC 1799
 1149 ----- 1148
 1800 AAGAGACAAGCCTACCTTCAAAAAAGAGAGATCAAGATCAAAAGATTAAGAGGACA 1859
 1149 ----- -AGATCAAAAGATGAAGAGATTAAGAGGACA 1178
 1860 AAGAAACCAAGAGAGAGTTAAGAGCTTCATGGATCGAAAGAGGATTTACAGAAAGTTA 1919
 1179 AAGAAACCAAGAGAGAGTTAAGAGCTTCATGGATCGAAAGAGGATTTACAGAAAGTTA 1238
 1920 AGTCGAGAGATGGAGAAATTCATGACCCCAAACTTAACATACACTAGAGATCTTCAGCC 1979
 1239 AGTCGAGAGATGGAGAAATTCATGACCCCAAACTTAACATACACTAGAGATCTTCAGCC 1298
 1980 GCCCTGAGGAGGAGGCGCAGCTGGACACCAAGAGGCTGAGGGCGCCGCCAGGTGGAG 2039
 1299 GCCCTGAGGAGGAGGCGCAGCTGGACACCAAGAGGCTGAGGGCGCCGCCAGGTGGAG 1358
 2040 CCGGCAAAAGGCTGGAGAGCTTCGTCTGCTGGGGGAGACCGAGAGCGAAGAGTTCCG 2099
 1359 CCGGCAAAAGGCTGGAGAGCTTCGTCTGCTGGGGGAGACCGAGAGCGAAGAGTTCCG 1418
 2100 AGAAGCTCAACACAGACAGCAGAGGAGGCGCTTCGAGCTGGAGAACTCAAGAAAAAGA 2159
 1419 AGAAGCTCAACACAGACAGCAGAGGAGGCGCTTCGAGCTGGAGAACTCAAGAAAAAGA 1478
 2160 GGGAGGAGAGAGAGAGTCTGGAGGAGGAGAGACAGAGAGAGAGAGAGAGAGAGCGG 2219
 1479 GGGAGGAGAGAGAGAGTCTGGAGGAGGAGAGACAGAGAGAGAGAGAGAGAGAGCGG 1538
 2220 ATCGAAACTCAGAGAGAGAGAGAGAGAGAGGCTTAAGAGAGAGAGTTGAAAGCGGAA 2279
 1539 ATCGAAACTCAGAGAGAGAGAGAGAGAGAGGCTTAAGAGAGAGAGTTGAAAGCGGAA 1598
 2280 GAGCAGAGAGCTGCTGAGAAAAAGCAGAGAGATGCCAGAGAGATCGCTTGTGATGACAAGA 2339
 1599 GAGCAGAGAGCTGCTGAGAAAAAGCAGAGAGATGCCAGAGAGATGCCAGAGAGAT 1658
 2340 AACCATTAAGTGTTCATCTCAAGTTCTCATCTCAAGATAGAGAGCGAGAGAT 2399
 1659 AACCATTAAGTGTTCATCTCAAGTTCTCATCTCAAGATAGAGAGCGAGAGAT 1718
 2400 TTTTGAATAAGTCTGTGAGAAAAAGCAGTGGTGTCAAATCGACCCCATCAAGCAGCAATAG 2459
 1719 TTTTGAATAAGTCTGTGAGAAAAAGCAGTGGTGTCAAATCGACCCCATCAAGCAGCAATAG 1778
 2460 TCTCCAGATTCAGACAGAGCTGGAGAGAGTATACAGTGCATTTAGGGACAGAAAAAGCG 2519
 1779 TCTCCAGATTCAGACAGAGCTGGAGAGAGTATACAGTGCATTTAGGGACAGAAAAAGCG 1838
 2520 CAAGACCTCAAGAGCGGCGAGCTCGATCTTCCTGCTGCTGAAGGTGTACGCAACA 2579
 1839 CAAGACCTCAAGAGCGGCGAGCTCGATCTTCCTGCTGAAGGTGTACGCAACA 1898
 2580 TCAAGAGTATGTGGAGAAAAAGGAAATGTGTTTTATCCGCCCATCGAGAGGAGGACACCAA 2639
 1899 TCAAGAGTATGTGGAGAAAAAGGAAATGTGTTTTATCCGCCCATCGAGAGGAGGACACCAA 1958
 2640 ATAGGAACTCTGGCTGAAGGTAGGGGTTTCTAGCGGCATCAATGAATGCTAACTA 2699
 1959 ATAGGAACTCTGGCTGAAGGTAGGGGTTTCTAGCGGCATCAATGAATGCTAACTA 2018

2700 AAACCCAGAGTGAACCAAGTCACTGCTCCCAACCTTCTGACTTGAGACCAGGAGCG 2759
 2019 AAACCCAGAGTGAACCAAGTCACTGCTCCCAACCTTCTGACTTGAGACCAGGAGCG 2078
 2760 TATCCAGCAGCGAAACCTCTGGGAAAAAGCAATCTGTGATAAGTCACTTCCCCCACTA 2819
 2079 TATCCAGCAGCGAAACCTCTGGGAAAAAGCAATCTGTGATAAGTCACTTCCCCCACTA 2138
 2820 AGGTTTGAAGCAGTTCAGAGAAAGCAAGCTCAAGACGAGGAGAGCTCAGTTGTAG 2879
 2139 AGGTTTGAAGCAGTTCAGAGAAAGCAAGCTCAAGACGAGGAGAGCTCAGTTGTAG 2198
 2880 AGGCTAAATTCGCTCTGTTTGTATTTATGTTGATTTACTAAATTTGGGTTCAATATCTTT 2939
 2199 AGGCTAAATTCGCTCTGTTTGTATTTATGTTGATTTACTAAATTTGGGTTCAATATCTTT 2258
 2940 TATTTTCAATATCCAGTAAGCAAGCTGATATTTACTATATTTAATTAATCAAGTCT 2999
 2259 TATTTTCAATATCCAGTAAGCAAGCTGATATTTACTATATTTAATTAATCAAGTCT 2318
 3000 AGAGATGTTTCATGTAAGTACTGCTTTTCACAGAGAGCCTGTTTCTAAAGAAACCCAT 3059
 2319 AGAGATGTTTCATGTAAGTACTGCTTTTCACAGAGAGCCTGTTTCTAAAGAAACCCAT 2378
 3060 GCTGTGAAATPAGAGACTTTTCTACTGATCATATACTCTGATCTGAGCAGTGATACCA 3119
 2379 GCTGTGAAATPAGAGACTTTTCTACTGATCATATACTCTGATCTGAGCAGTGATACCA 2438
 3120 ACCACATCTCAAGTCAACAGAGATCCAAAGTTTAAATTTGCTCGGGAATGTGTCAGTA 3179
 2439 ACCACATCTCAAGTCAACAGAGATCCAAAGTTTAAATTTGCTCGGGAATGTGTCAGTA 2498
 3180 TCTAGAAAAATGAACCGTAGTTTTTTTAAATPAGAGAGTCAAGTCTGTTCTGCA 3239
 2499 TCTAGAAAAATGAACCGTAGTTTTTTTAAATPAGAGAGTCAAGTCTGTTCTGCA 2558
 3240 CTTTATAATTAAGCATGGAAGAAATTAATCTAGTAGGCAATTTGAACACTTTTGAAGT 3299
 2559 CTTTATAATTAAGCATGGAAGAAATTAATCTAGTAGGCAATTTGAACACTTTTGAAGT 2618
 3300 AACCCATTTCAAGTTGAAATPAGTGGGATATGTTGCTTTAAAAAAGAGAAAGATG 3359
 2619 AACCCATTTCAAGTTGAAATPAGTGGGATATGTTGCTTTAAAAAAGAGAAAGATG 2678

RESULT 5
 AAQ28988
 ID AAQ28988 standard; cdna to mRNA; 2215 BP.

XX AAQ28988;

XX AC

XX 25-MAR-2003 (revised)

DT 25-FEB-1993 (first entry)

XX XX

XX Sequence encoding a polypeptide of human caldesmon designated type II (light) AA.

DE DE

XX Caldesmon; calmodulin; actin; binding; actomyosin system; regulation; ss.

XX Homo sapiens.

XX OS

XX Key

XX Location/Qualifiers

XX CDS

XX 12..1685

XX /*tag= a

XX polyA_site

XX 2099..2215

XX /*tag= c

XX polyA_signal

XX 2103..2108

XX /*tag= b

XX EP506377-A2.

XX 30-SEP-1992.

XX PD

XX XX

1899 TGTTCTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATCAACTCTG 1958
3101 TATCTGAGCAGTATACCAACACATCTGAAGTCAACAGAGATCCAGTTTAAATTCG 3160
1959 TATCTGAGCAGTATACCAACACATCTGAAGTCAACAGAGATCCAGTTTAAATTCG 2018
3161 CTGCGGAATGTGTCAGPATCTAGAAAAATGAACCGTAGTTTGTGTTTTTAAATACAG 3220
2019 CTGCGGAATGTGTCAGPATCTAGAAAAATGAACCGTAG- TTTTGTGTTTTTAAATACAG 2077
3221 AAGTCATGTTGTTCTGACATTTATAATAAGCATGGAAGAAATATCTTAGTAGCAAT 3280
2078 AAGTCATGTTGTTCTGACATTTATAATAAGCATGGAAGAAATATCTTAGTAGCAAT 2137
3281 TGTAAACACTTTTGAAGTAACCCATTTTCAAGATTTGAAATATCTGCGATATATGTTGCTTT 3340
2138 TGTAAACACTTTTGAAGTAACCCATTTTCAAGATTTGAAATATCTGCAATATATGTTGCTTT 2197
3341 AAAAAAAGAAAAAGA 3357
2198 TAAAAAAGAAAAA 2214
RESULT 6
AAH23102
ID AAH23102 standard; DNA; 3342 BP.
XX AC AAH23102;
XX
DT 17-SEP-2001 (first entry)
XX
DE Osteoarthritis tissue-derived nucleic acid sequence #32.
XX
XX Osteoarthritis; infectious disorder; Crohn's disease; sepsis; human;
KW wound healing; osteopathic; anti-arthritis; anti-inflammatory; vulnery;
KW antibacterial; antiallergic; ds.
XX
OS Homo sapiens.
XX WO200153531-A2.
XX
PD 26-JUL-2001.
XX
PF 18-JAN-2001; 2001WO-US000016.
XX
PR 18-JAN-2000; 2000US-0176523P.
XX
XX (PHAA) PHARMACIA CORP.
XX
XX Phippard D, Vasanthakamur G, Dotson S, Ma X;
XX WPI; 2001-451914/48.
XX
PT Substantially purified protein, polypeptide or their fragments, used to
PT identify a biologically active compound or composition and treat
PT mammalian osteoarthritis.
XX
XX
PS Claim 1; Page 118-119; 144pp; English.
XX
XX Sequences AAH23071-23152 represent nucleic acid sequences derived from
XX osteoarthritis tissues. The sequences are useful as probes and for the
XX diagnosis or prognosis of mammalian osteoarthritis. The polynucleotides
XX and polypeptides of the invention are useful for generating diagnostic
XX reagents, as targets for small molecule drug development, generation of
XX therapeutics, and cloning genes. Specific antibodies are used to generate
XX enzyme linked immunosorbent assays for detection of osteoarthritis. The
XX invented molecules can be used to treat osteoarthritis or to analyse the
XX disease-modifying activity of osteoarthritis drugs. Other disorders
XX treatable using the nucleic acid sequences include atopic, inflammatory
XX and infectious disorders e.g. Crohn's disease and sepsis, and wound
XX healing

SQ Sequence 3342 BP; 1108 A; 590 C; 687 G; 957 T; 0 U; 0 Other;
Query Match 39.3%; Score 1455.8; DB 4; Length 3342;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1482; Conservative 0; Mismatches 7; Indels 2; Gaps 2;
QY 1871 GAAGAAGTTAAGAGCTTCATGATCGAAAGAGGATTTACAGAGTTAAGTCGCGAAT 1930
Db 1 GAAGAAGTTAAGAGCTTCATGATCGAAAGAGGATTTACAGAGTTAAGTCGCGAAT 60
QY 1931 GGAGAAATTCATGACCCCAAACTTAAACATCTAGAGATCTTTAGCCGCCCTGGAGG 1990
Db 61 GGAGAAATTCATGACCCCAAACTTAAACATCTAGAGATCTTTAGCCGCCCTGGAGG 120
QY 1991 AGGCCGAGCGTGGACACCAAGAGGCTGAGGGCGCCCCCAGGTGAAGCGCGCAAAAG 2050
Db 121 AGGCCGAGCGTGGACACCAAGAGGCTGAGGGCGCCCCCAGGTGAAGCGCGCAAAAG 180
QY 2051 CTGGAGGAGCTTCGTCGTCGCGGGGAGACCGAGAGCGAAGAGTTTCGAGAGCTCAA 2110
Db 181 CTGGAGGAGCTTCGTCGTCGTCGCGGGGAGACCGAGAGCGAAGAGTTTCGAGAGCTCAA 240
QY 2111 CAGAGCAGCAGGAGGCGGCTTTGAGCTGGAGGAACTCAAGAAAAAGAGCGGAGGAGA 2170
Db 241 CAGAGCAGCAGGAGGCGGCTTTGAGCTGGAGGAACTCAAGAAAAAGAGCGGAGGAGA 300
QY 2171 AGGAAGTCTCTGGAGGAGGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2230
Db 301 AGGAAGTCTCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 2231 AGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2290
Db 361 AGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 2291 GCTGAGAAACGCCAGAGATGCCAGAGATGGCTTCTCAGATCAACAAGAACCAATCAAG 2350
Db 421 GCTGAGAAACGCCAGAGATGCCAGAGATGGCTTCTCAGATCAACAAGAACCAATCAAG 480
QY 2351 TGTTTCACTCTTAAAGGTTTCACTCTCAGATAGAGAGGCGGAGGAGGAGGAGGAGGAG 2410
Db 481 TGTTTCACTCTTAAAGGTTTCACTCTCAGATAGAGAGGCGGAGGAGGAGGAGGAGGAG 540
QY 2411 TCTGTGCAGAAAGCAGTGGTGTCAATCGACCCATCAAGCAGCAATAGTCTCAAGATT 2470
Db 541 TCTGTGCAGAAAGCAGTGGTGTCAATCGACCCATCAAGCAGCAATAGTCTCAAGATT 600
QY 2471 GACAGCAGCTGGAGCAGTATACAGTGCATTTGAGGAGCAAAAGCGGCAAAACCTTACA 2530
Db 601 GACAGCAGCTGGAGCAGTATACAGTGCATTTGAGGAGCAAAAGCGGCAAAACCTTACA 660
QY 2531 AAGCCGGCAGCTCGGATCTTCTGTCTGCTGAAGGTTAGCGCAACATCAAGAGTATG 2590
Db 661 AAGCCGGCAGCTCGGATCTTCTGTCTGCTGAAGGTTAGCGCAACATCAAGAGTATG 720
QY 2591 TGGGAGAAAGGAATGTGTTTTCATCCCACTGCGAGCAGCACCACCAATAGGAACCT 2650
Db 721 TGGGAGAAAGGAATGTGTTTTCATCCCACTGCGAGCAGCACCACCAATAGGAACCT 780
QY 2651 G-CTGCTTGAAGGTAGGGGTTTCTAGCCGATCAATGAATGGCTAATAAACCCCA 2709
Db 781 GCTGCTTGAAGGTAGGGGTTTCTAGCCGATCAATGAATGGCTAATAAACCCCA 840
QY 2710 TGGAAACAAAGTCACTGCTCCCAAACTTCTGAGTTCAGACGAGGAGAGCTATCAGCAA 2769
Db 841 TGGAAACAAAGTCACTGCTCCCAAACTTCTGAGTTCAGACGAGGAGAGCTATCAGCAA 900
QY 2770 GGGGACCTCTGGGAAAGCAATCTGTGATAGGTAC- TTTCCCACTAAGGTTTTCAG 2828
Db 901 GGGGACCTCTGGGAAAGCAATCTGTGATAGGTAC- TTTCCCACTAAGGTTTTCAG 960
QY 2829 ACAGTTCCAGAAAGAACCCCAAGCTCAGACGCGAGGAGCTAGTTTGTAGAGGCTTAT 2888
Db 961 ACAGTTCCAGAAAGAACCCCAAGCTCAGACGCGAGGAGCTAGTTTGTAGAGGCTTAT 1020

XX
PT polypeptide(s) with calmodulin- and actin binding activities - useful for
FT treating and diagnosing cancer and regulating digestive tract motility as
PT vasodilators.
XX
PS Disclosure; Page 31-32; 47pp; English.
XX
CC cDNA clones that encode polypeptide of human caldesmon were selected from
CC a cDNA library prep'd. from HeLa cells, and the DNA sequence and the amino
CC acid sequence encoding human caldesmon were deduced by DNA sequence
CC analysis. Next, the inventors prepared polypeptides of various lengths,
CC and looked for the functional domain of human caldesmon polypeptide,
CC using calmodulin-binding activity and actin-binding activity as indices.
CC There are two forms of caldesmon. The form of higher mol. wt. is called
CC type I and the form of lower mol. wt. is called type II. 312AA is a
CC polypeptide that is the C-terminal portion of types I and II. 122AA and
CC 118A are polypeptides that have a sequence of 116 AAs in the C-terminal
CC sequence of 312AA, and 94AA and 90AA are polypeptides that have a
CC sequence of 88 AAs that are in the central portion of 118AA, lacking
CC both the N-terminal portion and the C-terminal portion. On the N-terminal
CC of 118AA, there are residues Met-Ala that originated from a PCR primer,
CC and on the N-terminal of 90AA, there are residues Met-Ala that also
CC originated from a PCR primer. On the N-terminal of 122AA and 94AA, there
CC are residues Met-Ala that originated from PCR primer. The polypeptide
CC (typeI, typeII, 312AA, 122AA, 118AA, and 116AA) that contains the
CC polypeptide of 116 AAs has both calmodulin-binding activity and actin-
CC binding activity. Type I, type II, 312AA, 122AA, 118AA, 116AA all have
CC troponomyosin-binding activity and have an inhibitory activity to
CC atropomyosin ATPase, and so the functional unit of human caldesmon is
CC identified as being 116AA. (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
CC Sequence 2137 BP: 737 A: 424 C: 579 G: 397 T: 0 U: 0 Other;

Query Match	39.2%;	Score 1452.8;	DB 2;	Length 2137;				
Best local Similarity	95.3%;	Pred. No. 0;						
Matches 1530;	Conservative	0;	Mismatches	72;	Indels	4;	Gaps	3;
QY	1752	AGGAGAAAGAGAGGGAACTAAAGTCGACGCTAAAGAGAAAAGCTCCAAGNAGCAACG	1811					
DB	535	AAGAAAACAAGAAAAGAGCAAGAGAAAAGGAGGAGGAGGAGGAAGCAGAGAGGCCAAAGCGGAG	594					
QY	1812	CTACCTTCAAAAAAGAGAGATCAAAGATGAAAAGATTAAAAAGGACAAAGAACCCCAAAG	1871					
DB	595	GGAGCATTTGGAGAAAATCAGATCAAGATGAAAAGATTAAAAAGGAGAACCCAAAGCGAG	654					
QY	1872	AAGAAAGTTAAGAGCTTCATGGATCGAAAAAGAGGATTTACAGAGTTTAAGTCGCGAAGTG	1931					
DB	655	GGAGCATTTAAGAGCTTCATGGATCGAAAAGAGGGATTTACAGAAAGTTTAAGTCGCGAAGTG	714					
QY	1932	GGAATTCATGACCCCAAAACTTTAAACATACTGAGAATACTTTCAAGCCGCCCTCGAGGGGA	1991					
DB	715	GAGAATTCATGACCCCAAAACTTTAAACATACTGAGATACTTTCAAGCCGCCCTCGAGGGGA	774					
QY	1992	GGGCCAGCGTGGACACCAAGAGGCTGAGGGCGCCCGCCAGGTGGAAGCCGCGCAAAAGGC	2051					
DB	775	GGGCCAGCGTGGACACCAAGAGGCTGAGGGCGCCCGCCAGGTGGAAGCCGCGCAAAAGGC	834					
QY	2052	TGAGAGAGCTTCGTCTGCTCGCGGGAGACCCGAGAGCGGAGAGTTTCGAGAAAGCTCAAAAC	2111					
DB	835	TGAGAGAGCTTCGTCTGCTCGCGGGAGACCCGAGAGCGGAGAGTTTCGAGAAAGCTCAAAAC	894					
QY	2112	AGAAGCAGCAGAGGCGGCGCTTTTGAGGCTTGGAGGAACTCAAGAAAAAGAGGGAGGAGAGAA	2171					
DB	895	AGAAGCAGCAGAGGCGGCGCTTTTGAGGCTTGGAGGAACTCAAGAAAAAGAGGGAGGAGAGAA	954					
QY	2172	GGNAGTCTCGAGGAGGAAGACGAGAGGAGGAGACGAGGAAGCCGATCGAAAACTCA	2231					
DB	955	GGNAGTCTCGAGGAGGAAGACGAGAGGAGGAGGAGGAAGCCGATCGAAAACTCA	1014					
QY	2232	GAGAGGAGGAAGAGAAAGAGGAGGCTTAAAGGGAAGAGATTTAAAGGGCGAAGACGAGAAGCTG	2291					
DB	1015	CTACCTGAGAGGAGGAGGAGGCTTTAAGGGAAGAGATTTAAAGGGCGAAGACGAGCAGAGAGCTG	1074					

	Matches 1139;	Conservative	0;	Mismatches	88;	Indels	0;	Gaps	0;
QY	1601	CAGAAACGTAA	CAACAGCAGCTAG	AAGAGAGAAAAA	CAATGCCATCG	AGAGACAAGAGAT	AAAA	1660	
Db	451	CAAGAAAGATAT	CGAGATAGAGGAA	ACAGAAACAGT	CACCAAGTCCT	ACCAGAGAAT	GTAT	510	
QY	1661	GGGGAAAAGGT	AGAACACAGAAAAT	TAGAAAGGAAAT	TGGTAAATG	AAAAAGAAAGCA	AGAA	1720	
Db	511	TGGAGGGAT	GCTGAGANAACA	RAGAAAGACA	AGGANAAGGAGG	AGCAGGAAGAGG	AG	570	
QY	1721	GATAAACTT	CAGACAGCTGCT	TAAAGAAACAG	GGAGAGAGAGG	AACTTAAAGT	GC	1780	
Db	571	AAGCCAAAC	CGAGGAGCAAT	TGGAGAAATC	AGGGAGAGAGG	AGAACTTAAAGT	GC	630	
QY	1781	GCTAAAGAGAAA	AGCTTCCAAAGAAC	CAAGCCTACCTT	CAAAAAGAAAGAGAT	CAAAAGAT		1840	
Db	631	GCTAAAGAGAAA	AGCTTCCAAAGAAC	AGCCTTACCTT	CAAAAAGAAAGAGAT	CAAAAGAT		690	
QY	1841	GAAAAGATT	AAAAAGGACAAAG	NACCCAAAGAGAGT	TAAAGCTTCAT	CGATCG	GAAG	1900	
Db	691	GAAAAGATT	AAAAAGGACAAAG	NACCCAAAGAGAGT	TAAAGCTTCAT	CGATCG	GAAG	750	
QY	1901	AAGGATTTAC	AGAACTTAAAGT	CGCAGAAATTC	ATGACCCCAAA	ACTTAAAGAT		1960	
Db	751	AAGGATTTAC	AGAACTTAAAGT	CGCAGAAATTC	ATGACCCCAAA	ACTTAAAGAT		810	
QY	1961	ACTGAGAAAT	ACTTTACAGCCCT	TGAGGAGGCGC	CAGCTGGACAC	CAAGAGGCT	CAG	2020	
Db	811	ACTGAGAAAT	ACTTTACAGCCCT	TGAGGAGGCGC	CAGCTGGACAC	CAAGAGGCT	CAG	870	
QY	2021	GGCGCCCCC	CAGGTGGAAG	CGGCAAAAGGCT	TGGAGAGCTT	CGTCGTCGCGGG	AG	2080	
Db	871	GGCGCCCCC	CAGGTGGAAG	CGGCAAAAGGCT	TGGAGAGCTT	CGTCGTCGCGGG	AG	930	
QY	2081	ACCGAGAGG	AGAGTTCG	AGAGCTCAAACAG	AGCAGCAGGAGG	CGCTTTGAGCTG		2140	
Db	931	ACCGAGAGG	AGAGTTCG	AGAGCTCAAACAG	AGCAGCAGGAGG	CGCTTTGAGCTG		990	
QY	2141	GAGGAACCT	CAAGAAAAAG	GGGAGGAGAG	GAAGGTCTCTG	GAGGAAGAGCAGAGG		2200	
Db	991	GAGGAACCT	CAAGAAAAAG	GGGAGGAGAG	GAAGGTCTCTG	GAGGAAGAGCAGAGG		1050	
QY	2201	AGGAACGAGG	AGAACGCCGAT	CGAAACTCAG	AGAGGAGGAG	GAAGAGAGGAGGCT	TAAG	2260	
Db	1051	AGGAACGAGG	AGAACGCCGAT	CGAAACTCAG	AGAGGAGGAG	GAAGAGGAGGCT	TAAG	1110	
QY	2261	GAAGAGATT	GAAGGGCGAAG	AGCAGAAAGCT	CTGAGAAACGCC	AGAAATGC	CAGAAAT	2320	
Db	1111	GAAGAGATT	GAAGGGCGAAG	AGCAGAAAGCT	CTGAGAAACGCC	AGAAATGC	CAGAAAT	1170	
QY	2321	GGCTTCTCAG	ATCAGAAACCAAT	TCAAGTGTTTCACT	CTCTTAAAGGTTCACT	CTCTCAAG		2380	
Db	1171	GGCTTCTCAG	ATCAGAAACCAAT	TCAAGTGTTTCACT	CTCTTAAAGGTTCACT	CTCTCAAG		1230	
QY	2381	ATAGAAAGCG	AGCAAAATTTT	TGAAATGTGTG	CGAAGAGAGG	AGTGGTTC	CAATCG	2440	
Db	1231	ATAGAAAGCG	AGCAAAATTTT	TGAAATGTGTG	CGAAGAGAGG	AGTGGTTC	CAATCG	1290	
QY	2441	ACCCATCAAG	CAGCAATAGTCT	CCAAAGATTG	CACAGCAGACT	CGAGCAGATAT	ACCAGTGCA	2500	
Db	1291	ACCCATCAAG	CAGCAATAGTCT	CCAAAGATTG	CACAGCAGACT	CGAGCAGATAT	ACCAGTGCA	1350	
QY	2501	ATTGAGGGAA	CAAAAAAGCG	CAAAACCTT	CAAGCCGCGAG	CTTCGATCTT	CTGTTCTCT	2560	
Db	1351	ATTGAGGGAA	CAAAAAAGCG	CAAAACCTT	CAAGCCGCGAG	CTTCGATCTT	CTGTTCTCT	1410	
QY	2561	GCTGAAGGTG	PACGCAACAT	CAAGAGTATGT	GGGAGAAAGG	GAATGTGTTTT	CATCCCC	2620	
Db	1411	GCTGAAGGTG	PACGCAACAT	CAAGAGTATGT	GGGAGAAAGG	GAATGTGTTTT	CATCCCC	1470	
QY	2621	ACTGACAGG	CAACCCAAAT	TAAGGAAAT	CTGCTGGCTG	CAAGTAGGGT	TTCTAGCCG	2680	
Db	1471	ACTGACAGG	CAACCCAAAT	TAAGGAAAT	CTGCTGGCTG	CAAGTAGGGT	TTCTAGCCG	1530	

QY	2681	ATCAATGAATGGCTAACTAAAAACCCAGATGGAAACAAGTCACCTGCTCCCAACCTTCT	2740
Db	1531	ATCAATGAATGGCTAACTAAAAACCCAGATGGAAACAAGTCACCTGCTCCCAACCTTCT	1590
QY	2741	GACTTGAGACGAGGAGCGTATCAGCAAGCGGAACTCTGGGAAAGCAATCTGTGGAT	2800
Db	1591	GACTTGAGACGAGGAGCGTATCAGCAAGCGGAACTCTGGGAAAGCAATCTGTGGAT	1650
QY	2801	AAGGTCACTTCCCCCACTAAGGTTGA	2827
Db	1651	AAGGTCACTTCCCCCACTAAGGTTGA	1677
RESULT 9			
AAQ28985	AAQ28985 standard; cDNA to mRNA; 1599 BP.		
ID	AAQ28985		
XX	AC	AAQ28985;	
XX	AC		
XX	DT	25-MAR-2003 (revised)	
DT	25-FEB-1993	(first entry)	
XX	Sequence encoding a polypeptide of human caldesmon designated type II		
DE	DE	(light) AA.	
XX	KW	Caldesmon; calmodulin; actin; binding; actomyosin system; regulation; ss.	
XX	OS	Homo sapiens.	
XX	XX		
XX	Key	Location/Qualifiers	
FT	CDS	1..1599	
FT		/*tag= a	
XX	XX		
XX	PN	EP506377-A2.	
XX	PD	30-SEP-1992.	
XX	XX		
XX	PF	25-MAR-1992; 92EP-00302591.	
XX	XX		
XX	PR	29-MAR-1991; 91JP-00089106.	
XX	PR	27-DEC-1991; 91JP-00358040.	
XX	XX		
PA	(SOBU/)	SOBUE K.	
PA	(TAKI)	TAKARA SHUZO CO LTD.	
PI	Hayashi K, Asada K, Hashida T, Kotani H, Kato I, Sobue K;		
XX	XX		
DR	WPI; 1992-325651/40.		
DR	P-PSDB; AAR27362.		
XX	XX		
PT	Polypeptide(s) with calmodulin- and actin binding activities - useful for treating and diagnosing cancer and regulating digestive tract motility as vasodilators.		
PT	XX		
XX	XX		
PS	Claim 5; Page 27-28; 47pp; English.		
XX	XX		
CC	cDNA clones that encode polypeptide of human caldesmon were selected from a cDNA library prepd. from HeLa cells, and the DNA sequence and the amino acid sequence encoding human caldesmon were deduced by DNA sequence analysis. Next, the inventors prepared polypeptides of various lengths, and looked for the functional domain of human caldesmon polypeptide, using calmodulin-binding activity and actin-binding activity as indices. There are two forms of caldesmon. The form of higher mol. wt. is called type I and the form of lower mol. wt. is called type II. 312AA is a polypeptide that is the C-terminal portion of types I and II. 122AA and 118A are polypeptides that have a sequence of 116 AAs in the C-terminal sequence of 312AA, and 94AA and 90AA are polypeptides that have a sequence of 88 AAs that are in the central portion of 118AA, lacking both the N-terminal portion and the C-terminal portion. On the N-terminal of 118AA, there are residues Met-Ala that originated from a PCR primer, and on the N-terminal of 90AA, there are residues Met-Ala that also originated from a PCR primer. On the N-terminal of 122AA and 94AA, there		

1364 GGGGAAAGGGAATGTGTTTTTCATCCCCCACTGCAGCGGCACACCAAAATAGGAAACATG 1424

2652 CTGGCTTGAAGGTAGGGTTTCTAGCCGATCAATGAATGGGTAACTAAAACCCCGAGATG 2711

1424 CTGGCTTGAAGGTAGGGTTTCTAGCCGATCAATGAATGGGTAACTAAAACCCCGAGATG 1483

2712 GAAACAATGACCTGCTCTCCAAACCTTCTGACTTGAGACACGAGAGACGTTATCCAGCAAGC 2771

1484 GAAACAATGACCTGCTCTCCAAACCTTCTGACTTGAGACACGAGAGACGTTATCCAGCAAGC 1543

2772 GGAACCTCTGGGAAAGCAATCTGTGGATAAGGTCACTTCCCCCACTAAGGTTTGA 2827

1544 GGAACCTCTGGGAAAGCAATCTGTGGATAAGGTCACTTCCCCCACTAAGGTTTGA 1599

RESULT 10
AAQ28984
ID AAQ28984 standard; cDNA to mRNA; 939 BP.
XX AC AAQ28984;
XX AC
XX AC
XX AC
DT 25-MAR-2003 (revised)
DT 25-FEB-1993 (first entry)
DE Sequence encoding a polypeptide of human caldesmon designated 312 AA.
XX Caldesmon; calmodulin; actin; binding; actomyosin system; regulation; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 1. 939
FT CDS /*tag= a
FT FT
XX EP506377-A2.
XX
XX 30-SEP-1992.
XX
XX 25-MAR-1992; 92EP-00302591.
XX
XX 29-MAR-1991; 91JP-00089106.
XX 27-DEC-1991; 91JP-00358040.
XX
XX (SOBU/) SOBUE K.
XX (TAKI) TAKARA SHUZO CO LTD.
XX
XX Hayaashi K, Asada K, Hashida T, Kotani H, Kato I, Sobue K;
XX
XX WPI; 1992-325651/40.
XX P-PSDB; AAR27361.
XX
XX Polypeptide(s) with calmodulin- and actin binding activities - useful for
XX treating and diagnosing cancer and regulating digestive tract motility as
XX vasodilators.
XX
XX Claim 5; Page 26; 47pp; English.
XX
XX cDNA clones that encode polypeptide of human caldesmon were selected from
XX a cDNA library prepd. from HeLa cells, and the DNA sequence and the amino
XX acid sequence encoding human caldesmon were deduced by DNA sequence
XX analysis. Next, the inventors prepared polypeptides of various lengths,
XX and looked for the functional domain of human caldesmon polypeptide,
XX using calmodulin-binding activity and actin- binding activity as indices.
XX There are two forms of caldesmon. The form of higher mol. wt. is called
XX type I and the form of lower mol. wt. is called type II. 312AA is a
XX polypeptide that is the C-terminal portion of types I and II. 122AA and
XX 118A are polypeptides that have a sequence of 116 AAs in the C-terminal
XX sequence of 312AA, and 94AA and 90AA are polypeptides that have a
XX sequence of 88 AAs that are in the central portion of 118AA, lacking
XX both the N-terminal portion and the C-terminal portion. On the N-terminal
XX of 118AA, there are residues Met-Ala that originated from a PCR primer,
XX and on the N-terminal of 90AA, there are residues Met-Ala that also
XX originated from a PCR primer. On the N-terminal of 122AA and 94AA, there

CC are residues Met-Ala that originated from PCR primer. The polypeptide
CC (typeI, typeII, 312AA, 122AA, 118AA, and 116AA) that contains the
CC polypeptide of 116 AAs has both calmodulin-binding activity and actin-
CC binding activity. Type I, type II, 312AA, 122AA, 118AA, 116AA all have
CC tropomyosin-binding activity and have an inhibitory activity to
CC actomyosin ATPase, and so the functional unit of human caldesmon is
CC identified as being 116AA. (Updated on 25-MAR-2003 to correct PN field.)
XX (Updated on 25-MAR-2003 to correct PA field.)
SQ Sequence 939 BP; 314 A; 194 C; 275 G; 156 T; 0 U; 0 Other;

Query Match 25.2%; Score 932.6; DB 2; Length 939;
Best Local Similarity 99.6%; Pred. No. 5.5e-208;
Matches 935; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1889 ATGGATCGAAGAGGAGTTTACAGAAGTTAAGTCGCAGAAATGAGAAATCATGACCCAC 1948
Db 1 ATGGATCGAAGAGGAGTTTACAGAAGTTAAGTCGCAGAAATGAGAAATCATGACCCAC 60

QY 1949 AAATTAACATACTAGAAATATCTTCAGCGCCCTGGAGGAGCCAGGTCGACACC 2008
Db 61 AAATTAACATACTAGAAATATCTTCAGCGCCCTGGAGGAGCCAGGTCGACACC 120

QY 2009 AAGGAGGCTGAGGCGCCGCCAGGTCGAAAGCCGCAAAAGGCTGGAGGAGCTTCGTGCT 2068
Db 121 AAGGAGGCTGAGGCGCCGCCAGGTCGAAAGCCGCAAAAGGCTGGAGGAGCTTCGTGCT 180

QY 2069 CFTGCGGGAGACCGAGAGCGAAGAGTTGAGAGCTCAACAGAGCAGCAGGAGCG 2128
Db 181 CFTGCGGGAGACCGAGAGCGAAGAGTTGAGAGCTCAACAGAGCAGCAGGAGCG 240

QY 2129 GCTTTGAGCTGAGGAGCTCAAGAAAGAGGAGGAGAGAGAGTCTCTGGAGGAG 2188
Db 241 GCTTTGAGCTGAGGAGCTCAAGAAAGAGGAGGAGAGAGTCTCTGGAGGAG 300

QY 2189 GAAGACAGAGGAGAGAGCGAGGAGCGGATCGAATACTCAAGAGGAGGAGAGAG 2248
Db 301 GAAGACAGAGGAGAGAGCGAGGAGCGGATCGAATACTCAAGAGGAGGAGAGAG 360

QY 2249 AGAGGCTAAGGAGAGAGATTGAAGCGGAGAGAGAGAGTCTGAGAAAGCCAGAG 2308
Db 361 AGAGGCTAAGGAGAGAGATTGAAGCGGAGAGAGAGTCTGAGAAAGCCAGAG 420

QY 2309 ATGCCAGAGATGCTTGTGAGATGACAGAAACCAATCAAGTGTTCATCTTAAGT 2368
Db 421 ATGCCAGAGATGCTTGTGAGATGACAGAAACCAATCAAGTGTTCATCTTAAGT 480

QY 2369 TCATCTCTCAAGTAGAGAGCGAGCAATTTTGAATAAGTCTGTGAGAAAGCAGT 2428
Db 481 TCATCTCTCAAGTAGAGAGCGAGCAATTTTGAATAAGTCTGTGAGAAAGCAGT 540

QY 2429 GGTGTCAAATCGACCCATCAAGCAGCAATAGTCTCAAGATTGACAGAGACTGGAGCG 2488
Db 541 GGTGTCAAATCGACCCATCAAGCAGCAATAGTCTCAAGATTGACAGAGACTGGAGCG 600

QY 2489 TATACCAAGTCAATGAGGAGCAAAAGCCGCAAACTTACAAAGCCGAGCTCGGAT 2548
Db 601 TATACCAAGTCAATGAGGAGCAAAAGCCGCAAACTTACAAAGCCGAGCTCGGAT 660

QY 2549 CTTCTCTGTTCTGTTGAAGGTTGACGCAACATCAAGAGTATGTGGAGAAAGGAATGTG 2608
Db 661 CTTCTCTGTTCTGTTGAAGGTTGACGCAACATCAAGAGTATGTGGAGAAAGGAATGTG 720

QY 2609 TTTTCATCCCCACCTGAGCAGGACACCAATAGGAAGTCTGCTGTTGAGGTAGGG 2668
Db 721 TTTTCATCCCCACCTGAGCAGGACACCAATAGGAAGTCTGCTGTTGAGGTAGGG 780

QY 2669 GTTTCAGCCGATCAATGAATGGTGTAACTAAACCCAGATGGAACCAAGTCACTGTG 2728
Db 781 GTTTCAGCCGATCAATGAATGGTGTAACTAAACCCAGATGGAACCAAGTCACTGTG 840

QY 2729 CCCAAACCTTCTGATTTGAGACAGAGAGCGTATCCAGAGGCGAACCTCTGGGAAAG 2788

Db 841 CCCAAACCTTCTGATTTGAGACAGAGAGCTATCCAGCAAGCGAACCTCTGGGAAAG 900

QY 2789 CAATCTGTGGATAAGTCACTTCCCTCCCTCACTAAGGTTTGA 2827
Db 901 CAATCTGTGGATAAGTCACTTCCCTCCCTCACTAAGGTTTGA 939

RESULT 11
AAC00905
ID AAC00905 standard; cDNA; 705 BP.
XX AAC00905;
AC AAC00905;
DT 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 903.
DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX EP1033401-A2.
XX 06-SEP-2000.
XX 21-FEB-2000; 2000EP-00200610.
XX 26-FEB-1999; 99US-0122487P.
XX (GEST) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX P-PSDB; AAG00899.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX Claim 1; SEQ ID NO 903; 71pp + Sequence Listing; English.
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the
XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX derived from 30 different tissues. EST sequences usually correspond
XX mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX well suited for isolating cDNA sequences derived from the 5' ends of
XX mRNAs and even in those cases where longer cDNA sequences have been
XX obtained the full 5' UTR is rarely included. 5' ESTs are derived from
XX mRNAs with intact 5' ends and can therefore be used to obtain full length
XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX gene therapy and chromosome mapping procedures. They are used to obtain
XX upstream regulatory sequences and to design expression and secretion
XX vectors
SQ Sequence 705 BP; 190 A; 171 C; 200 G; 143 T; 0 U; 1 Other;

Query Match 18.7%; Score 691; DB 3; Length 705;
Best Local Similarity 99.9%; Pred. No. 1.8e-151;
Matches 691; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCTAGGAATGACAGGAGTCTCCACAGGAGGTCATCCACCTTGGTGGGTGTCGT 60
Db 14 GCCTAGGAATGACAGGAGTCTCCACAGGAGGTCATCCACCTTGGTGGGTGTCGT 73

QY 61 CATTGGTGGTCTATTAGAAAACGACAGAGCAATGCATACCAACCGCTCCCGACTGTAAA 120
Db 74 CATTGGTGGTCTATTAGAAAACGACAGAGCAATGCATACCAACCGCTCCCGACTGTAAA 133

QY 121 CATAGGGATATGTGTTTCATCTAGCATGGACTTCTGGGAGGGGCCAAGGAAGCGGCTCT 180

134 CATAGGGGATATGTGTTACCTAGCATGACATCTTGGGAGGGGCCCAAGAAAGGCGGTCT 193
181 GGAGTTTATTAGTAGAGAGTGTATTTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
194 GGAGTTTATTAGTAGAGAGTGTATTTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 253
241 CTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
254 CTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 313
301 AGTTTCAGACTACAGAGTATCATTTGGAACATTTCAAGATCATCAATCAATTTCCACAGG 360
314 AGTTTCAGACTACAGAGTATCATTTGGAACATTTCAAGATCATCAATCAATTTCCACAGG 373
361 GATTGGTGACCAACAGAGGCTCAGATCATCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTG 420
374 GATTGGTGACCAACAGAGGCTCAGATCATCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTG 433
421 GTCTCCCTGAACTGAAATCACACCATGATGATTTTGAAGCTGCGAGAGAACTTAGAAG 480
434 GTCTCCCTGAACTGAAATCACACCATGATGATTTTGAAGCTGCGAGAGAACTTAGAAG 493
481 GCAAAAGAGGAGGAGATGCGACTCGAAGCAGAAAGATGCGCTACACAGAGGAATGACGA 540
494 GCAAAAGAGGAGGAGATGCGACTCGAAGCAGAAAGATGCGCTACACAGAGGAATGACGA 553
541 TGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
554 TGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 613
601 GAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
614 GAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 673
661 CAGTGTGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 692
674 CAGTGTGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 705

RESULT 12
AA39869 standard; DNA; 756 BP.
AA39869;
02-JUL-1999 (first entry)
Gastric cancer associated gene.
Cancer associated antigen; diagnosis; research; treatment; human;
breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
prostate cancer; ss.
Homo sapiens.
WO9904265-A2.
28-JAN-1999.
15-JUL-1998; 98WC-US014679.
17-JUL-1997; 97US-00896164.
10-OCT-1997; 97US-0061599P.
10-OCT-1997; 97US-0061765P.
10-OCT-1997; 97US-00948705.
11-OCT-1997; 97GB-00021697.
22-JUN-1998; 98US-00102322.
(LUDW-) LUDWIG INST CANCER RES.
Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;
O'hare M, Obata Y, Pfreundschuh M, Tureci O, Sahin U;

XX WPI; 1999-132448/11.
XX New isolated cancer associated nucleic acids and polypeptides - isolated
PT using sera from cancer patients, used to develop products for the
PT diagnosis, monitoring or treatment of cancers.
XX Claim 67; Page 578; 787pp; English.
XX The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer.
XX Sequence 756 BP; 316 A; 117 C; 244 G; 77 T; 0 U; 2 Other;
SQ Query Match 18.6%; Score 689.2; DB 2; Length 756;
Best Local Similarity 97.0%; Pred. No. 4.9e-151;
Matches 733; Conservative 0; Mismatches 20; Indels 3; Gaps 3;
QY 1500 CAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1559
DB 1 CAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60
QY 1560 GGGCCAGGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1619
DB 61 GGGCCAGGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 1620 TAGAAG 1679
DB 121 TAGAAG 180
QY 1680 AAATAGAGAGGAG 1739
DB 181 AAATAGAGAGGAG 240
QY 1740 TCCTAAAG 1799
DB 241 TCCTAAAG 300
QY 1800 AAGAAG 1859
DB 301 AAGAAG 360
QY 1860 AAGAAG 1919
DB 361 AAGAAG 420
QY 1920 AGTCGAG 1979
DB 421 AGTCGAG 480
QY 1980 GCCCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2039
DB 481 GCCCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
QY 2040 CCGGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2098
DB 541 CCGGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 2099 GAGAAGCTCAAAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2157
DB 601 GAGAAGCTCAAAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660

Db	433	ACATCTGAGTCAACAGAGATCCAGTTTAAATTCCTCGGGAATGTGTGAGTAICT	492
Qy	3183	AGAAAAATGACCGTAGTTTGTGTTTAAATACAGAGTCACTGTTCTGCACTT	3242
Db	493	AGAAAAATGACCGTAGTTTGTGTTTAAATACAGAGTCACTGTTCTGCACTT	552
Qy	3243	TATAATAAGCATGGAAGAAATTAATCTTAGTAGGCAATTTGTAACATTTTGAAGTAAC	3302
Db	553	TATAATAAGCATGGAAGAAATTAATCTTAGTAGGCAATTTGTAACATTTTGAAGTAAC	612
Qy	3303	CCATTTTCAGATTTGAAATCTCGGTAATGTTTCTTTAAAAAAGAAAGATG	3359
Db	613	CCATTTTCAGATTTGAAATCTCGGTAATGTTTCTTTAAAAAAGAAAGATG	669
RESULT 15			
AAH33117	ID	AAH33117 standard; cDNA; 2511 BP.	
XX	AC	AAH33117;	
XX	DT	03-SEP-2001 (first entry)	
XX	DE	Human colon cancer antigen encoding cDNA SEQ ID NO:173.	
XX	DE	Human; colon cancer; colon cancer antigen; diagnosis; detection;	
XX	KW	colorectal carcinoma; chromosome 7; ss.	
XX	OS	Homo sapiens.	
XX	PN	WO200122920-A2.	
XX	PD	05-APR-2001.	
XX	PF	28-SEP-2000; 2000WO-US026524.	
XX	PR	29-SEP-1999; 99US-0157137P.	
XX	PR	03-NOV-1999; 99US-0163280P.	
XX	PA	(HUMA-) HUMAN GENOME SCI INC.	
XX	PI	Ruben SM, Barash SC, Birse CE, Rosen CA;	
XX	DR	WPI: 2001-235357/24.	
XX	DR	P-PSDB; AAG73866.	
XX	PT	Nucleic acids encoding 4277 human colon cancer-associated polypeptides, useful for preventing, diagnosing and/or treating colorectal cancers.	
XX	PS	Claim 1; Page 2329-2330; 9803pp; English.	
XX	CC	AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytostatic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing inactive proteins or to supplement the patient's own production of P. Additionally, N may be used to produce the colon cancer-associated Ps, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent sequences used in the exemplification of the present invention. N.B. Pages 666 to 682 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for SEQ ID NO:1027 to 1052, 7921 and 7922	
XX	XX	Sequence 2511 BP; 814 A; 445 C; 443 G; 803 T; 0 U; 6 Other;	
XX	XX	Query Match	
XX	XX	15.8%; Score 584.2; DB 4; Length 2511;	

		Best Local Similarity 98.7%; Pred. No. 3.3e-126;				Matches 589; Conservative 0; Mismatches 8; Indels 0; Gaps 0;			
QY	2763	CCAGCAAGCGGAACCTCTGGGAAAGCAATCTGTGATAAGGTCACTTCCCCCACTAAGG	2822						
Db	73	CGAGCAAGCGGAACCTCTGGGAAAGCAATCTGTGATAAGGTCACTTCCCCCACTAAGG	132						
QY	2823	TTTGAGACAGTTCAGAAAGAACCCAGCTCAAGCGGAGGAGCTCAGTTGTAGAGG	2882						
Db	133	TTTGAGACAGTTCAGAAAGAACCCAGCTCAAGCGGAGGAGCTCAGTTGTAGAGG	192						
QY	2883	GCTAATTCGCTCTGTTTGTATTTATGTTGATTTACTAAATTTGGGTTCATTATCTTTTAT	2942						
Db	193	GCTAATTCGCTCTGTTTGTATTTATGTTGATTTACTAAATTTGGGTTCATTATCTTTTAT	252						
QY	2943	TTTTCAATATCCCAAGTAAACCCAGTATATATCACTATATTTAATATATCAAGTCTAGA	3002						
Db	253	TTTTCAATATCCCAAGTAAACCCAGTATATATCACTATATTTAATATATCAAGTCTAGA	312						
QY	3003	GATGTTTCATGTAAGTAAAGTACTGCTTTTCACAGGAGCCTGTTTCTAAAGAAACCCATGCT	3062						
Db	313	GATGTTTCATGTAAGTAAAGTACTGCTTTTCACAGGAGCCTGTTTCTAAAGAAACCCATGCT	372						
QY	3063	GTGAAATAGAGACTTTTCTACTGATCATCAACTCTGTATCTGAGCAGTATACCAACC	3122						
Db	373	GTGAAATAGAGACTTTTCTACTGATCATCAACTCTGTATCTGAGCAGTATACCAACC	432						
QY	3123	ACATCTGAAGTCAACAGAGATCCCAAGTTTAAATTCCTCGGGAATGTGCGAGTATCT	3182						
Db	433	ACATCTGAAGTCAACAGAGATCCCAAGTTTAAATTCCTCGGGAATGTGCGAGTATCT	492						
QY	3183	AGAAAAATGAACCGTAGTTTTTTTGTATTTTAAATACAGAAATCATGTTGTTCTGCACIT	3242						
Db	493	AGAAAAATGAACCGTAGTTTTTTTGTATTTTAAATACAGAAATCATGTTGTTCTGCACIT	552						
QY	3243	TATAATAAGCATGCAAGAAATTTCTTAGTAGGCAATTTGTACACTTTTGAAGTAAC	3302						
Db	553	TATAATAAGCATGCAAGAAATTTCTTAGTAGGCAATTTGTACACTTTTGAAGTAAC	612						
QY	3303	CCATTTTCAGATTTGAAATACCTGCGATAATGGTTGCTTTAAAAAAGAAAAAGATG	3359						
Db	613	CCATTTTCAGATTTGAAATACCTGCGATAATGGTTGCTTTAAAAAAGAAAAAGATG	669						

Search completed: March 10, 2004, 11:00:07
Job time : 916.994 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

DM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:19:42 ; Search time 169.055 Seconds
(without alignments)
12152.440 Million cell updates/sec

Title: US-10-084-817-30

Perfect score: 3702

Sequence: 1 gcctagggaatgacaggcat.....aaggggagggtaccgaggaac 3702

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1570.2	42.4	2215	1	US-08-285-440-14
2	1570.2	42.4	2215	1	US-08-630-349-14
3	1481.6	40.0	2137	1	US-08-285-440-13
4	1481.6	40.0	2137	1	US-08-630-349-13
5	1087.8	29.4	1677	1	US-08-285-440-12
6	1087.8	29.4	1677	1	US-08-630-349-12
7	999.2	27.0	1599	1	US-08-285-440-11
8	999.2	27.0	1599	1	US-08-630-349-11
9	937.4	25.3	939	1	US-08-285-440-10
10	937.4	25.3	939	1	US-08-630-349-10
11	428.6	11.6	606	3	US-09-385-982-295
12	350.8	9.5	369	1	US-08-285-440-9
13	350.8	9.5	369	1	US-08-630-349-9
14	350.4	9.5	357	1	US-08-285-440-8
15	350.4	9.5	357	1	US-08-630-349-8
16	346.4	9.4	348	1	US-08-285-440-7
17	346.4	9.4	348	1	US-08-630-349-7
18	264.2	7.1	273	1	US-08-285-440-26
19	263	7.1	285	1	US-08-630-349-26
20	263	7.1	285	1	US-08-285-440-23
21	263	7.1	285	1	US-08-630-349-23
22	173	4.7	7218	1	US-08-232-463-14
23	137.4	3.7	53526	3	US-08-658-136-2
24	137.4	3.7	53577	3	US-08-658-136-1
25	116.8	3.2	3489	2	US-08-728-323A-1
26	116.8	3.2	3489	4	US-09-298-568-1
27	116.8	3.2	3489	4	US-09-410-399-1

c 28	116.8	3.2	32207	2	US-08-770-379-20	Sequence 20, Appl
c 29	116.8	3.2	32207	3	US-08-757-669A-20	Sequence 20, Appl
c 30	116.8	3.2	32207	4	US-09-230-371A-20	Sequence 20, Appl
c 31	114.2	3.1	5433	3	US-08-929-329-1	Sequence 1, Appl
c 32	112.6	3.0	6755	3	US-08-931-999-4	Sequence 4, Appl
c 33	102.8	2.8	1926	4	US-09-249-585A-2	Sequence 2, Appl
c 34	102.8	2.8	1926	4	US-09-410-399-3	Sequence 3, Appl
c 35	102.8	2.8	2580	3	US-09-050-863-2	Sequence 2, Appl
c 36	102.8	2.8	2580	4	US-09-359-081-2	Sequence 2, Appl
c 37	102.8	2.8	5452	2	US-09-130-114-1	Sequence 1, Appl
c 38	102.8	2.8	8705	4	US-09-647-344A-14	Sequence 14, Appl
c 39	102.8	2.8	9600	3	US-08-910-647-1	Sequence 1, Appl
c 40	102.8	2.8	9600	4	US-09-620-925-1	Sequence 15, Appl
c 41	102.8	2.8	10596	1	US-07-884-811-15	Sequence 1, Appl
c 42	102.8	2.8	10596	1	US-07-885-971-15	Sequence 15, Appl
c 43	102.8	2.8	10596	1	US-08-087-783A-15	Sequence 15, Appl
c 44	102.8	2.8	10596	1	US-08-194-088B-15	Sequence 15, Appl
c 45	102.8	2.8	10596	2	US-08-194-087-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-285-440-14
; Sequence 14, Application US/08285440
; Patent No. 5532337
; GENERAL INFORMATION:
; APPLICANT: Ken'ichiro HAYASHI et al.
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/285,440
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/858,947
; FILING DATE: March 27, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2215 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORGANISM: Human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:

3101 TATCTGACAGTGATACCAACACATCTGAAGTCAACAGAGATCCAAAGTTTAAATTC 3160
Db TATCTGACAGTGATACCAACACATCTGAAGTCAACAGAGATCCAAAGTTTAAATTC 2018
3161 CTGGGAATGTGTCAGTATCTAGAAAATGAACCGTAGTTTGTGTTTAAATACAG 3220
Db CTGGGAATGTGTCAGTATCTAGAAAATGAACCGTAG-TTTTGTGTTTAAATACAG 2077
3221 AAGTCATGTGTTCTGCACTTATATATAAAGCATGGAAGAAATATCTTAGTAGGCAAT 3280
Db AAGTCATGTGTTCTGCACTTATATATAAAGCATGGAAGAAATATCTTAGTAGGCAAT 2137
3281 TGTAACATTTTGAAGTAACCATTTTCAGATTTGAAATACGCGATAATGTTGCTTT 3340
Db TGTAACATTTTGAAGTAACCATTTTCAGATTTGAAATACGCGATAATGTTGCTTT 2197
3341 AAAAAAAGAAAAAGA 3357
Db TAAAAAAGAAAAAGA 2214

RESULT 2

US-08-630-349-14
; Sequence 14, Application US/08630349
; Patent No. 5739008
; GENERAL INFORMATION:
; APPLICANT: Ken'ichiro HAYASHI et al.
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,349
; FILING DATE: April 10, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/285,440
; FILING DATE: August 4, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/858,947
; FILING DATE: March 27, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2215 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE: Human
; IMMEDIATE SOURCE: HeLa Cell
; FEATURE:
; NAME/KEY:
; LOCATION: 12-1695
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "human caldesmon of higher
; OTHER INFORMATION: molecular weight"
; FEATURE:
; NAME/KEY:
; LOCATION: 2103-2108
; IDENTIFICATION METHOD:

; OTHER INFORMATION: /note= "poly A signal"
; FEATURE:
; NAME/KEY:
; LOCATION: 2099-2215
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "poly A site"
US-08-630-349-14
Query Match 42.4%; Score 1570.2; DB 1; Length 2215;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 93; Indels 4; Gaps 3;
QY 1601 CAGAAACGTAAACAGCAGCTAGAAAGAGAAAAACATGCCATGCAAGAGACAAAGATAAAA 1660
Db 462 CAAGAAAGATACAGATACAGGAAACAGAAACAGTCACCAAGTCTCTACCAGAGAAATGAT 521
QY 1661 GGGGAAAAGGTAGAACAGAAAATAGAGGGAATGGTAAATGAAAGAAAGACCAAGAA 1720
Db 522 TGGAGGGATGCTGAAAGAAAACAAAGAAAGACAGAGGAAAAGGAGGAGAGAGAGAG 581
QY 1721 GATAAACTTCAGACAGCTGCTCTTAAAGAAACAGGAGAGAGAGAGGAACTTAAAGTGCAA 1780
Db 582 AAGCCAAAGCGAGGAGCATTGCGAGAAATTCAGGAGAGAGAGAGGAACTTAAAGTGCAA 641
QY 1781 GCTAAAAGAGAAAAGCTCCAAAGAGCAAGCCCTACCTTCAAAAAGAGAGATCAAAGAT 1840
Db 642 GCTAAAAGAGAAAAGCTCCAAAGAGCAAGCCCTACCTTCAAAAAGAGAGATCAAAGAT 701
QY 1841 GAAAAGATTAAAAGAGCAAAAGAACCCAAAGAAAGAGTTAAGAGCTTCATGGATCGAAAG 1900
Db 702 GAAAAGATTAAAAGAGCAAAAGAACCCAAAGAAAGAGTTAAGAGCTTCATGGATCGAAAG 761
QY 1901 AAGGATTTACAGAACTTAAGTCGCGAATGAGAGAAATCATGACCCCAAAACTTAAACAT 1960
Db 762 AAGGATTTACAGAACTTAAGTCGCGAATGAGAGAAATCATGACCCCAAAACTTAAACAT 821
QY 1961 ACTGAGAACTATTTTCAGCGCCCTCGAGGAGAGGCGCCAGCGTGGACACCAAGAGAGCTGAG 2020
Db 822 ACTGAGAACTATTTTCAGCGCCCTCGAGGAGAGGCGCCAGCGTGGACACCAAGAGAGCTGAG 881
QY 2021 GCGGCCCCCAGGTGAAAGCGGCAAAAGGCTGAGAGAGCTTCGTGTCGCGGGGAG 2080
Db 882 GCGGCCCCCAGGTGAAAGCGGCAAAAGGCTGAGAGAGCTTCGTGTCGCGGGGAG 941
QY 2081 ACCGAGAGCGAAGAGTTCCGAGAACTCAACAGAGCAGAGGAGGCGGCTTTGGAGCTG 2140
Db 942 ACCGAGAGCGAAGAGTTCCGAGAACTCAACAGAGCAGAGGAGGCGGCTTTGGAGCTG 1001
QY 2141 GAGGAACTCAAGAAAAGAGGAGAGAGAGAAAGGCTCTTGGAGAGGAGAGAGAGAGG 2200
Db 1002 GAGGAACTCAAGAAAAGAGGAGAGAGAGAAAGGCTCTTGGAGAGGAGAGAGAGAGG 1061
QY 2201 AGGAAAGAGGAG 2260
Db 1062 AGGAAAGAGGAG 1121
QY 2261 GAAGAGATTGAAAGGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2320
Db 1122 GAAGAGATTGAAAGGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1181
QY 2321 GCCTTGTGATGACAGAAAG 2380
Db 1182 GCCTTGTGATGACAGAAAG 1241
QY 2381 ATAGAAAGAGCGAGCAGAAATTTTGAATTAAGTCTGTGAGAAAAGCAGTGGTGTCAATCG 2440
Db 1242 ATAGAAAGAGCGAGCAGAAATTTTGAATTAAGTCTGTGAGAAAAGCAGTGGTGTCAATCG 1301
QY 2441 ACCCATCAAGCAGCAATAGTCTCCAGATTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 2500
Db 1302 ACCCATCAAGCAGCAATAGTCTCCAGATTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 1361
QY 2501 ATTGAGGGAACAAAAAGCGCAAAACCTTACAAAGCGGAGGCTTCGGATCTTCTGTCTCT 2560

Db	1362	ATTGAGGGAAACAAAAGCGCCAAACCTACAAGCGCGAGCCTCGGATCTTCCCTGTTCCCT	1421
Qy	2561	GCTGAAGGTGTACGCAACATCAAGAGTATGTGGGAGAAAGGGAATGTGTTTTCATCCCCC	2620
Db	1422	GCTGAAGGTGTACGCAACATCAAGAGTATGTGGGAGAAAGGGAATGTGTTTTCATCCCCC	1481
Qy	2621	ACTGCAGGAGGACACACCAATAAAGGAACCTGCTGGCTTGAAGGTAGGGTTCCTAGAGCGC	2680
Db	1482	ACTGCAGGAGGACACACCAATAAAGGAACCTGCTGGCTTGAAGGTAGGGTTCCTAGAGCGT	1541
Qy	2681	ATCAATGAATGGCTAACTAAAAACCCAGATGGAAACAAGTCACCTGTCCCAACCTTCT	2740
Db	1542	ATCAATGAATGGCTAACTAAAAACCCAGATGGAAACAAGTCACCTGTCCCAACCTTCT	1601
Qy	2741	GACTTGAGACGAGGAGACGTATCCAGCAAGCGGAACCTCTGGGAAGAACAATCTGTGGAT	2800
Db	1602	GACTTGAGACGAGGAGACGTATCCAGCAAGCGGAACCTCTGGGAAGAACAATCTGTGGAT	1661
Qy	2801	AAGGTCACCTCCCCCACTAAGGTTTGAGACAGTTCCAGAAGAACCACCAAGCTCAAGAGCG	2860
Db	1662	AAGGTCACCTCCCCCACTAAGGTTTGAGACAGTTCCAGAAGAACCACCAAGCTCAAGAGCG	1721
Qy	2861	AGGACGAGCTCAGTTGTAGAGGCGTAATTCGCTCTGTGTTGCTATTATGTTGATTACTA	2920
Db	1722	AGGACGAGCTCAGTTGTAGAGGCGTAATTCGCTCTGTGTTGCTATTATGTTGATTACTA	1779
Qy	2921	AAATGGGTTCAATTATCTTTTATTTTCCAAATCCAGTAAACCCATGTATATTATCACTA	2980
Db	1780	AAATGGGTTCAATTATCTTTTATTTTCAATATCCAGTAAACCCATGTATATTATCACTA	1839
Qy	2981	TATTTAAATAACACAGCTAGAGATGTTCAATGTTAAAGTACTGCCTTGCACAGAGCC	3040
Db	1840	TATTTAAATAACACAG - CTAGAGATGTTCAATGTTAAAGTACTGCCTTGCACAGAGCC	1898
Qy	3041	TGTTTCTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATATAACTCTG	3100
Db	1899	TGTTTCTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATATAACTCTG	1958
Qy	3101	TATCTGAGCAGTGATACCAACCAATCTGAAGTCAAAGAGATCCAGTTTAAATTCG	3160
Db	1959	TATCTGAGCAGTGATACCAACCAATCTGAAGTCAAAGAGATCCAGTTTAAATTCG	2018
Qy	3161	CTGCGGAATGTGTCAGTACTAGAAAAATGAAACCGTAGTTTCTGTTTTTTAAATACAG	3220
Db	2019	CTGCGGAATGTGTCAGTACTAGAAAAATGAAACCGTAG - TTTTGTTTTTTAATACAG	2077
Qy	3221	AAGTCATGTTGTTCTGCATTTATAATAAGCATGGAAGAAATTACTTAGTAGGCAAT	3280
Db	2078	AAGTCATGTTGTTCTGCATTTATAATAAGCATGGAAGAAATTACTTAGTAGGCAAT	2137
Qy	3281	TGTAACACATTTTTGAAAGTAGTAAACCATTTTCAGATTGAAATCTGCGATAATGGTGTCTT	3340
Db	2138	TGTAACACATTTTTGAAAGTAGTAAACCATTTTCAGATTGAAATCTGCGATAATGGTGTCTT	2197
Qy	3341	AAAAAAAAAGAAAAAGA	3357
Db	2198	AAAAAAAAAAAAAAAAAA	2214

COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/285,440
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/858,947
FILING DATE: March 27, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 2137 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE: Human
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE: Hela Cell
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION: 12-1607
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "human caldesmon of lower
OTHER INFORMATION: molecular weight"
FEATURE:
NAME/KEY:
LOCATION: 2025-2030
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "poly A signal"
FEATURE:
NAME/KEY:
LOCATION: 2121-2137
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "poly A site"
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:

Db 1351 ATTGAGGGAACAAAAGCGCAAAACCTACAAAGCGGAGCCTCGGATCTTCCTGTTCT 1410
Qy 2561 GCTGAAGGTGTACCGCAACATCAAGAGTATGTGGGAGAAAGGGAATGTGTTTTCATCCCC 2620
Db 1411 GCTGAAGGTGTACCGCAACATCAAGAGTATGTGGGAGAAAGGGAATGTGTTTTCATCCCC 1470
Qy 2621 ACTGAGAGGACACCAATAAGGAACCTGCTGCTTGAAGGTAGGGTTTCTAGCCGC 2680
Db 1471 ACTGAGAGGACACCAATAAGGAACCTGCTGCTTGAAGGTAGGGTTTCTAGCCGC 1530
Qy 2681 ATCAATGAATGGCTAACTAAAAACCCAGATGGAACCAAGTCACCTGCTCCCAAACTTCT 2740
Db 1531 ATCAATGAATGGCTAACTAAAAACCCAGATGGAACCAAGTCACCTGCTCCCAAACTTCT 1590
Qy 2741 GACTTGAACACGAGACGTATCCAGAACGCGAACCTCTGGAAAGCAATCTGTGAT 2800
Db 1591 GACTTGAACACGAGACGTATCCAGAACGCGAACCTCTGGAAAGCAATCTGTGAT 1650
Qy 2801 AAGGTCACTTCCCCCACTAAGGTTGA 2827
Db 1651 AAGGTCACTTCCCCCACTAAGGTTGA 1677

RESULT 6

US-08-630-349-12
; Sequence 12, Application US/08630349
; Patent No 573008
; GENERAL INFORMATION:
; APPLICANT: Ken'ichiro HAYASHI et al.
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,349
; FILING DATE: April 10, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/285,440
; FILING DATE: August 4, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/858,947
; FILING DATE: March 27, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1677 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE: Human
; IMMEDIATE SOURCE: Hela Cell
; US-08-630-349-12

Query Match 29.4%; Score 1087.6; DB 1; Length 1677;
Best Local Similarity 92.9%; Pred. No. 1.7e-255;
Matches 1140; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
Qy 1601 CAGAAACGTAAACAGCAGCTAGAGAGAAACATGCGCATGCAGAGACAAAGATAAA 1660

Db 451 CAAGAAAGATACAGATAGAGGAAACAGAAACAGTCCACCAAGTCTTACCAGAGAAATGAT 510
Qy 1661 GGGGAAAAGGTAGAACAGAAAATAGAGGGAATGGTAAATGAAAGAAAGAACCAAGNA 1720
Db 511 TGGAGGGATGCTGAAGAAAACAAGAAAAGAGACAAGGAAAAGGAGGAGGAGAGGAG 570
Qy 1721 GATTAACCTTCAGACAGCTGCTTAAAGAAAACAGGAGAGAGAGAGGAACTTAAAGTGCAA 1780
Db 571 AAGCCAAAGCGGAGGAGCATTCGAGAAAATCAAGGAGAGAGAGGGAACATAAGTGCAA 630
Qy 1781 GTTAAAGAGAAAAGCTTCAAGAAAGCAAGCTTACCTTCAAAAAAAGAGAGATCAAGAT 1840
Db 631 GTTAAAGAGAAAAGCTTCAAGAAAGCAAGCTTACCTTCAAAAAAAGAGAGATCAAGAT 690
Qy 1841 GAAAGATTAATAAGAGAACCAAGAACCCCAAGAGAAAGTTAAGAGCTTCAATGATCGAAAG 1900
Db 691 GAAAGATTAATAAGAGAACCCCAAGAGAAAGTTAAGAGCTTCAATGATCGAAAG 750
Qy 1901 AAGGGATTTACAGAAAGTTAAGTCGAGAAATGGAGAAATTCATGACCCCAAACTTAAACAT 1960
Db 751 AAGGGATTTACAGAAAGTTAAGTCGAGAAATGGAGAAATTCATGACCCCAAACTTAAACAT 810
Qy 1961 ACTGAGATTAATTTAGCCGCTTGGAGGAGGCGCCAGCTGGACACCAAGAGGCTGAG 2020
Db 811 ACTGAGATTAATTTAGCCGCTTGGAGGAGGCGCCAGCTGGACACCAAGAGGCTGAG 870
Qy 2021 GGCGCCCCCAGGTGGAAGCGGCAAAAGGCTGGAGGAGCTTCGTCTGCTCGCGGGAG 2080
Db 871 GGCGCCCCCAGGTGGAAGCGGCAAAAGGCTGGAGGAGCTTCGTCTGCTCGCGGGAG 930
Qy 2081 ACCGAGGCGAAGATTCGAGAGCTCAACAGAGCAGCAGGAGGCGGCTTTGGAGCTG 2140
Db 931 ACCGAGGCGAAGATTCGAGAGCTCAACAGAGCAGCAGGAGGCGGCTTTGGAGCTG 990
Qy 2141 GAGGAATCAAGAAAAGAGGAGAGAGAGAGAGTCTCTGGAGGAGAAAGAGCAGAGG 2200
Db 991 GAGGAATCAAGAAAAGAGGAGAGAGAGAGAGTCTCTGGAGGAGAAAGAGCAGAGG 1050
Qy 2201 AGGAAGAGGAGAGGAGGAGTCAAGAACTCAGAGAGGAGGAGAGAGAGGAGGCTTAAAG 2260
Db 1051 AGGAAGAGGAGGAGGAGGAGTCAAGAACTCAGAGAGGAGGAGAGAGAGGAGGCTTAAAG 1110
Qy 2261 GAAGAGATTCGAAGCGGAGAGCAGAGAGCTGCTGAGAAACGCCAGAGATGCCAGAGAT 2320
Db 1111 GAAGAGATTCGAAGCGGAGAGCAGAGAGCTGCTGAGAAACGCCAGAGATGCCAGAGAT 1170
Qy 2321 GGCTTGTCAAGTGAACAAACCATTCAGGTGTTTCACTCTTAAAGGTTCACTCTCAAG 2380
Db 1171 GGCTTGTCAAGTGAACAAACCATTCAGGTGTTTCACTCTTAAAGGTTCACTCTCAAG 1230
Qy 2381 ATAGAGAGGAGGAGAGATTTTGAATAAGTCTGTGCAGAAAAGCAGTGGTGTCAATCG 2440
Db 1231 ATAGAGAGGAGGAGAGATTTTGAATAAGTCTGTGCAGAAAAGCAGTGGTGTCAATCG 1290
Qy 2441 ACCCATCAAGCAGCAATAGTCTCAAGATTCAGAGAGAGAGAGAGTATACAGTGCA 2500
Db 1291 ACCCATCAAGCAGCAATAGTCTCAAGATTCAGAGAGAGAGAGAGTATACAGTGCA 1350
Qy 2501 ATTGAGGAGCAAAAGCGCAAAACCTCAAGCGGAGCGCTCGGATCTTCTGTTCTCT 2560
Db 1351 ATTGAGGAGCAAAAGCGCAAAACCTCAAGCGGAGCGCTCGGATCTTCTGTTCTCT 1410
Qy 2561 GCTGAAGGTGTACGCAACATCAAGAGTATGTGGGAGAAAGGGAATGTGTTTCACTCCCC 2620
Db 1411 GCTGAAGGTGTACGCAACATCAAGAGTATGTGGGAGAAAGGGAATGTGTTTCACTCCCC 1470
Qy 2621 ACTGAGAGGACACCAACCAATAGGAAATGCTGCTTGAAGGTAGGGTTTCTAGCCGC 2680
Db 1471 ACTGAGAGGACACCAACCAATAGGAAATGCTGCTTGAAGGTAGGGTTTCTAGCCGC 1530
Qy 2681 ATCAATGAATGGCTAACTAAAAACCCAGATGGAACCAAGTCACCTCTCCCAAACTTCT 2740

1531 ATCAATGAATGGCTAACTATAAAACCCAGATGGAAACAAAGTCACCTGCTGCCAACCTCT 1590

2741 GACTTGAGACCAAGGAGACGATATCCAGCAACCGAAACCTCTGGGAAAAGCAATCTGTGGAT 2800

1591 GACTTGAGACCAAGGAGACGATATCCAGCAACCGAAACCTCTGGGAAAAGCAATCTGTGGAT 1650

2801 AAGGTCACCTTCCCCCACTAAGGTTTGA 2827

1651 AAGGTCACCTTCCCCCACTAAGGTTTGA 1677

RESULT 7

JS-08-285-440-11
Sequence 11, Application US/08285440
Patent No. 5532337
GENERAL INFORMATION:
APPLICANT: Ken'ichiro HAYASHI et al.
TITLE OF INVENTION: POLYPEPTIDE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESS: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/285.440
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/858,947
FILING DATE: March 27, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1599 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE: Human
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE: HeLa Cell
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:

Sequence 10, Application US/08285440
 Patent No. 5532337
 GENERAL INFORMATION:
 APPLICANT: Ken'ichiro HAYASHI et al.
 TITLE OF INVENTION: POLYPEPTIDE
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wenderoth, Lind & Ponack
 STREET: 805 Fifteenth Street, N.W., #700
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: DisplayWrite
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/285,440
 FILING DATE: 07/08/1992
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/858,947
 FILING DATE: March 27, 1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Warren M. Cheek, Jr.
 REGISTRATION NUMBER: 33,367
 REFERENCE/DOCKET NUMBER:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-8850
 TELEFAX:
 TELEX:
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 939 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL:
 ANTI-SENSE:
 FRAGMENT TYPE:
 ORIGINAL SOURCE: Human
 ORGANISM:
 STRAIN:
 INDIVIDUAL ISOLATE:
 DEVELOPMENTAL STAGE:
 HAPLOTYPE:
 TISSUE TYPE:
 CELL TYPE:
 CELL LINE:
 ORGANELLER:
 IMMEDIATE SOURCE: HeLa Cell
 LIBRARY:
 CLONE:
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT:
 MAP POSITION:
 UNITS:
 FEATURE:
 NAME/KEY:
 LOCATION:
 IDENTIFICATION METHOD:
 OTHER INFORMATION:
 PUBLICATION INFORMATION:
 AUTHORS:
 TITLE:
 JOURNAL:
 VOLUME:
 ISSUE:
 PAGES:
 DATE:

DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO:
 US-08-285-440-10
 Query Match 25.3%; Score 937.4; DB 1; Length 939;
 Best Local Similarity 99.9%; Pred. No. 6.1e-219;
 Matches 938; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1889 ATGGATCGAAAGAGGATTTACAGAAAGTTAACTCGCAGAAATGAGAAATTCATGACCCAC 1948
 DB 1 ATGGATCGAAAGAGGATTTACAGAAAGTTAACTCGCAGAAATGAGAAATTCATGACCCAC 60
 QY 1949 AAACCTAAACATACCTAGAAATCTTTACGCGCCCTCGAGAGGGCCAGCGTGGACACC 2008
 DB 61 AAACCTAAACATACCTAGAAATCTTTACGCGCCCTCGAGAGGGCCAGCGTGGACACC 120
 QY 2009 AAGGAGGCTGAGGGGCGCCCGCCAGGTGGAAGCGGCAAAAGGCTGGAGGAGCTTCGTCGT 2068
 DB 121 AAGGAGGCTGAGGGGCGCCCGCCAGGTGGAAGCGGCAAAAGGCTGGAGGAGCTTCGTCGT 180
 QY 2069 CGTCGCGGGAGACCGAGAGCGAAGATTTCGAGAAAGCTCAAAACAGAGCAGCAGAGGCG 2128
 DB 181 CGTCGCGGGAGACCGAGAGCGAAGATTTCGAGAAAGCTCAAAACAGAGCAGCAGAGGCG 240
 QY 2129 GCTTTGGAGCTGAGGAACTCAAGAAAGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAG 2188
 DB 241 GCTTTGGAGCTGAGGAACTCAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
 QY 2189 GAAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2248
 DB 301 GAAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
 QY 2249 AGGAGGCTAAAGGAGAGAGATTGAAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2308
 DB 361 AGGAGGCTAAAGGAGAGAGATTGAAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 QY 2309 ATGCCAGAGAGATGGCTTGTTCAGATGACAAAGAACCCATTCAAGTGTTCCTCTCTAAAGGT 2368
 DB 421 ATGCCAGAGAGATGGCTTGTTCAGATGACAAAGAACCCATTCAAGTGTTCCTCTCTAAAGGT 480
 QY 2369 TCATCTCTCAAGATAGAGAGCGAGAGAGATTTTGAATAGTCTGTCAGAAAGAGCAGT 2428
 DB 481 TCATCTCTCAAGATAGAGAGCGAGAGAGATTTTGAATAGTCTGTCAGAAAGAGCAGT 540
 QY 2429 GGTGTCAATCGAACCCATCAAGCAGCAATAGTCTCAAGATTGACAGCAGACTGGAGCAG 2488
 DB 541 GGTGTCAATCGAACCCATCAAGCAGCAATAGTCTCAAGATTGACAGCAGACTGGAGCAG 600
 QY 2489 TATACAGTGCAATTGAGGGAACAAAAGCGCAAAACCTCAAAAGCGGCGAGCCTCGGAT 2548
 DB 601 TATACAGTGCAATTGAGGGAACAAAAGCGCAAAACCTCAAAAGCGGCGAGCCTCGGAT 660
 QY 2549 CTTCTGTTCTGCTGTAAGGTGTACCAACATCAAGAGATGTGCGGAGAAAAGGGAATGTG 2608
 DB 661 CTTCTGTTCTGCTGTAAGGTGTACCAACATCAAGAGATGTGCGGAGAAAAGGGAATGTG 720
 QY 2609 TTTTCATCCCCCAGCTGAGCAGGCAACCAAAATAAGGAAACTGTGGTTGAAGTAGGG 2668
 DB 721 TTTTCATCCCCCAGCTGAGCAGGCAACCAAAATAAGGAAACTGTGGTTGAAGTAGGG 780
 QY 2669 GTTCTAGCGGCATCAATGATGGCTAACTTAAACCCAGATGGAACCAAGTCACCTGCT 2728
 DB 781 GTTCTAGCGGCATCAATGATGGCTAACTTAAACCCAGATGGAACCAAGTCACCTGCT 840
 QY 2729 CCAAAACCTTCTGACTTGAGACCGAGAGAGAGTATCCAGAGCGGAACTCTTGGGAAAAG 2788
 DB 841 CCAAAACCTTCTGACTTGAGACCGAGAGAGAGTATCCAGAGCGGAACTCTTGGGAAAAG 900
 QY 2789 CAATCTGTGGATAAGGTCACTTCCCCCAGTAAAGTTGA 2827
 DB 901 CAATCTGTGGATAAGGTCACTTCCCCCAGTAAAGTTGA 939

	QY	2249	AGGAGCGCTAAAGGAAGACAGATTCCAAAGGCCAAGCAGACAGAAGCTCTCTCAGAAAACGCCCAAG	2309
	DB	361	AGAGGCTTAAGGAAGAGATTTGAAGGCCGAGAGCAGAGCTGTCTGAGAAACCCCAAG	420
	QY	2309	ATGCCAGAAGATGGCTTGTGCAGATGACAAGAAACCAATTCOAAGTGTTTTCACTCTCAAAGGT	2368
	DB	421	ATGCCAGAAGATGGCTTGTGCAGATGACAAGAAACCAATTCOAAGTGTTTTCACTCTCAAAGGT	480
	QY	2369	TCACTCTCAAGATAGAAGCGAGCGAGCAGAAATTTTTGAAATAAGTCTGTGCGAAGAACGCACT	2428
	DB	481	TCACTCTCAAGATAGAAGCGAGCGAGCAGAAATTTTTGAAATAAGTCTGTGCGAAGAACGCACT	540
	QY	2429	GGTGTCAAATCGACCATTCAAGCAGCAATAGTCTCCAAGATTCGACAGCAGACTGGAGCAG	2488
	DB	541	GGTGTCAAATCGACCATTCAAGCAGCAATAGTCTCCAAGATTCGACAGCAGACTGGAGCAG	600
	QY	2489	TATACCAAGTGCAATTCAGGCGNACAAAAGCGCAAAACCTTACAAAGCGCGCAGCCTCGCAT	2548
	DB	601	TATACCAAGTGCAATTCAGGCGNACAAAAGCGCAAAACCTTACAAAGCGCGCAGCCTCGCAT	660
	QY	2549	CTTCTGTTTCTCTGCTGAAAGGTGTACCAACATCAAGAGTAGTGTGGGAGAAAGGGAATGTG	2608
	DB	661	CTTCTGTTTCTCTGCTGAAAGGTGTACCAACATCAAGAGTAGTGTGGGAGAAAGGGAATGTG	720
	QY	2609	TTTTTCATCCCCCACTGCAGCAGGCGCACCAAAATAAGGAAACTGCTGGCTTGAAGGTAGGG	2668
	DB	721	TTTTTCATCCCCCACTGCAGCAGGCGCACCAAAATAAGGAAACTGCTGGCTTGAAGGTAGGG	780
	QY	2669	GTTTCTAGCGCGCATCAATGAATGGCTAACTAAAAACCCCAGATGAAAACAAGTCACCTGCT	2728
	DB	781	GTTTCTAGCGCGCATCAATGAATGGCTAACTAAAAACCCCAGATGAAAACAAGTCACCTGCT	840
	QY	2729	CCCAACACCTTCTGACCTTGAGACCAAGAGCGTATCCAGCAAGCGGAACCTCTCGGGAAG	2788
	DB	841	CCCAACACCTTCTGACCTTGAGACCAAGAGCGTATCCAGCAAGCGGAACCTCTCGGGAAG	900
	QY	2789	CAATCTGTGGATAAGGTCACCTTCCCCCACTAAGGTTTTGA	2827
	DB	901	CAATCTGTGGATAAGGTCACCTTCCCCCACTAAGGTTTTGA	939
 RESULT 11				
US-09-385-982-295/c				
; Sequence 295, Application US/09385982				
; Patent No. 6262334				
; GENERAL INFORMATION:				
; APPLICANT: ENDEGE, WILSON O., ET AL.				
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION				
; TITLE OF INVENTION: PRODUCTS: 11				
; FILE REFERENCE: CCDNA-260XX				
; CURRENT APPLICATION NUMBER: US/09/385,982				
; CURRENT FILING DATE: 1999-08-30				
; EARLIER APPLICATION NUMBER: 09/328,111				
; EARLIER FILING DATE: 1999-06-08				
; EARLIER APPLICATION NUMBER: 60/117,393				
; EARLIER FILING DATE: 1999-01-27				
; EARLIER APPLICATION NUMBER: 60/098,639				
; EARLIER FILING DATE: 1998-08-31				
; NUMBER OF SEQ ID NOS: 544				
; SOFTWARE: FastSeq for Windows Version 3.0				
; SEQ ID NO 295				
; LENGTH: 606				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
; FEATURE:				
; NAME/KEY: misc.feature				
; LOCATION: (1)...(606)				
; OTHER INFORMATION: n = A,T,C or G				
US-09-385-982-295				

Query Match 11.6%; Score 428.6; DB 3; Length 606;
Best Local Similarity 96.5%; Pred. No. 6.1e-95;

```

Matches 437; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
2Y 2570 GTAGCAACATCAAGAGATGTGGGAGAAAGGAATGTTTCATCCCCCACTGCAGCA 2629
Db 457 GTAGCAACATCAAGAGATGTGGGAGAAAGGAATGTTTCATCCCCCACTGCAGCA 398
2Y 2630 GGCACACCAATAAGGAACTCTGGCTTGAAGGTAGGGGTTTCTAGCCGCATCAATGAA 2689
Db 397 GGCACACCAATAAGGAACTCTGGCTTGAAGGTAGGGGTTTCTAGCCGCATCAATGAA 338
2Y 2690 TGGCTAATCAAAACCCAGATGGAACAGTCACTCTCCCAACCTCTGACTTGAGA 2749
Db 337 TGGCTAATCAAAACCCAGATGGAACAGTCACTCTCCCAACCTCTGACTTGAGA 278
2Y 2750 CCAGGAGACGTATCCAGCAAGCGGACCTCTGGGAAAGCAATCTGTGGATAAGGTCACT 2809
Db 277 CCAGGAGACGTATCCAGCAAGCGGACCTCTGGGAAAGCAATCTGTGGATAAGGTCACT 218
2Y 2810 TCCCCCACTAAGTTTGACACAGTTCAGAAAGACCCAGCTCAAGACGACGACGAGC 2869
Db 217 TCCCCCACTAAGTTTGACACAGTTCAGAAAGACCCAGCTCAAGACGACGACGAGC 158
2Y 2870 TCAGTTGTAGAGGGGTAATTCGCTCTGTTTGTATTATTTGATTGTTTAAATTTGGTT 2929
Db 157 TCAGTTGTAGAGGGGTAATTCGCTCTGTTTGTATTATTTGATTGTTTAAATTTGGTT 98
2Y 2930 CATATCTTTTATTTTCAATATCCAGTAAGCCATGATATATCACTATATTTAATA 2989
Db 97 CATATCTTTTATTTTCAATATCCAGTAAGCCATGATATATCACTATATTTAATA 38
2Y 2990 ATCAGCTCTAGAGATGTTTCATGGTAAAGTAC 3022
Db 37 ATCAGCTCTAGAGATGTTTCATGGTAAAGTAC 5

```

RESULT 12

```

US-08-285-440-9
; Sequence 9, Application US/08285440
; Patent No. 5832337
; GENERAL INFORMATION:
; APPLICANT: Ken'ichiro HAYASHI et al.
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/285,440
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/858,947
; FILING DATE: March 27, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369 base pairs

```

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEITICAL:
; ANTI-SENSE:
; FRAGMENT TYPE: Human
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE: Hela Cell
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-08-285-440-9
; Query Match 9.5%; Score 350.8; DB 1; Length 369;
; Best Local Similarity 99.4%; Pred. No. 4.4e-76;
; Matches 352; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2474 AGCAGACTGGAGCAGTATACCACTGCAATTGAGGGAAACAAAAGCGCAAAACCTCAAG 2533
Db 16 AGCAGACTGGAGCAGTATACCACTGCAATTGAGGGAAACAAAAGCGCAAAACCTCAAG 75
QY 2534 CCGGCGAGCCTCGGATCTTCTGTTCTGCTGAAGGTGTACGCAACATCAAGAGTATGTGG 2593
Db 76 CCGGCGAGCCTCGGATCTTCTGTTCTGCTGAAGGTGTACGCAACATCAAGAGTATGTGG 135
QY 2594 GAGAAAGGGAATGTGTTTTCATCCCACTGCGAGCGACACCAAAATAGGAAACTGCT 2653
Db 136 GAGAAAGGGAATGTGTTTTCATCCCACTGCGAGCGACACCAAAATAGGAAACTGCT 195
QY 2654 GGCTTGAAGTAGGGGTTTCTAGCCGATCAATGAATGGCTAACTAAACCCCAATGGA 2713
Db 196 GGCTTGAAGTAGGGGTTTCTAGCCGATCAATGAATGGCTAACTAAACCCCAATGGA 255
QY 2714' AACAACTCACCTGCTCCCAACCTTCTGACTTGTAGACGAGAGAGTATCCAGCAAGCGG 2773
Db 256 AACAACTCACCTGCTCCCAACCTTCTGACTTGTAGACGAGAGAGTATCCAGCAAGCGG 315
QY 2774 AACCTCTGGGAAAGCAATCTGTGATAAGGTCACTTCCCACTAAAGTTTGA 2827
Db 316 AACCTCTGGGAAAGCAATCTGTGATAAGGTCACTTCCCACTAAAGTTTGA 369

```

RESULT 13

US-08-630-349-9
; Sequence 9, Application US/08630349
; Patent No. 5739008
; GENERAL INFORMATION:
; APPLICANT: Ken'ichiro HAYASHI et al.
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,349
; FILING DATE: April 10, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/285,440
; FILING DATE: August 4, 1994
; PRIOR APPLICATION NUMBER: 07/858,947
; FILING DATE: March 27, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; ORIGINAL SOURCE: Human
; IMMEDIATE SOURCE: Hela Cell
US-08-630-349-9
Query Match 9.5%; Score 350.8; DB 1; Length 369;
Best Local Similarity 99.4%; Pred. No. 4.4e-76;
Matches 352; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2474 AGCAGACTGGAGCAGTATACAGTGCATTTGAGGGAACAAAAAGCGCAAAACCTACAAAG 2533
Db 16 AGGAGACTGGAGCAGTATACAGTGCATTTGAGGGAACAAAAAGCGCAAAACCTACAAAG 75
QY 2534 CCGGAGCCTCGGATCTTCTGCTGAGGTGTACGCAACATCAAGATATGTGG 2593
Db 76 CCGGAGCCTCGGATCTTCTGCTGAGGTGTACGCAACATCAAGATATGTGG 135
QY 2594 GAGAAAGGGAATGCTTTTCATCCCACTGCAGCAGGACACCAATAAGGAAACTGCT 2653
Db 136 GAGAAAGGGAATGCTTTTCATCCCACTGCAGCAGGACACCAATAAGGAAACTGCT 195
QY 2654 GGCCTTGAAGGTAGGGGTTTCTAGCCGCATCAATGAATGGCTAACTAAACCCAGATGGA 2713
Db 196 GGCCTTGAAGGTAGGGGTTTCTAGCCGCATCAATGAATGGCTAACTAAACCCAGATGGA 255
QY 2714 AACAGTCACTGCTCCCAACCTTCTGACTTGAGACCGAGGACGATATCCAGCAAGCGG 2773
Db 256 AACAGTCACTGCTCCCAACCTTCTGACTTGAGACCGAGGACGATATCCAGCAAGCGG 315
QY 2774 AACCTCTGGGAAAGCAATCTGTGGATAAGGTCACTTCCCCCACTAAGGTTTGA 2827
Db 316 AACCTCTGGGAAAGCAATCTGTGGATAAGGTCACTTCCCCCACTAAGGTTTGA 369

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2004, 03:37:21 ; Search time 862.637 Seconds
(without alignments)
15799.545 Million cell updates/sec

Title: US-10-084-817-30

Perfect score: 3702

Sequence: 1 gctagggaatgacaggcat.....aaggggcggttacggagaac 3702

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata1/pubpna/US09A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata1/pubpna/US50_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3702	100.0	3702	14	US-10-084-817-30
2	3201.2	86.5	3500	15	US-10-062-674-1815
3	3116.4	84.2	3610	9	US-09-880-107-2406
4	3116.4	84.2	3610	10	US-09-873-267C-721
5	2353	63.6	4586	15	US-10-062-674-1817
6	1793.6	48.4	4547	9	US-09-974-298-36
7	1793.6	48.4	4547	14	US-10-084-817-31
8	1771	47.8	4261	15	US-10-062-674-1816
9	1578	42.6	2845	15	US-10-341-434-136
10	1578	42.6	2845	15	US-10-341-434-215
11	1455.8	39.3	3342	9	US-09-765-231A-32
12	584.2	15.8	2511	9	US-09-925-301-100
13	584.2	15.8	2511	14	US-10-106-698-183
14	546.8	14.8	2209	9	US-09-974-298-35
15	499	13.5	541	12	US-10-085-783A-16172

16	499	13.5	541	15	US-10-242-535A-16172
17	495.8	13.4	543	14	US-10-198-846-9612
18	447.4	12.1	449	14	US-10-060-036-1194
19	436.6	11.8	503	10	US-09-918-995-19285
20	428.6	11.6	436	10	US-09-918-995-3028
21	428.6	11.6	606	10	US-09-871-161-295
22	412.8	11.2	2850	14	US-10-198-846-13694
23	284	7.7	232	9	US-09-234-093B-5212
24	229.6	6.2	339	9	US-09-925-301-689
25	228.4	6.2	432	15	US-10-264-049-1570
26	211.4	5.7	400	9	US-09-960-352-3498
27	186.4	5.0	480	12	US-10-085-783A-50761
28	186.4	5.0	480	15	US-10-242-535A-50761
29	165.4	4.5	720	13	US-10-101-487-76
30	164.2	4.4	1840	14	US-10-017-161-2359
31	164.2	4.4	1840	15	US-10-292-798-2003
32	164	4.4	720	13	US-10-101-487-74
33	148.6	4.0	169	12	US-10-085-783A-18790
34	148.6	4.0	169	15	US-10-242-535A-18790
35	145	3.9	533	10	US-09-918-624B-55
36	137.4	3.7	53522	10	US-09-904-968A-1
37	135.2	3.7	554	13	US-10-101-487-106
38	133	3.6	554	13	US-10-101-487-69
39	132.8	3.6	530	13	US-10-101-487-73
40	132.6	3.6	522	13	US-10-101-487-71
41	129.6	3.5	68233	15	US-10-034-650-31
42	124.8	3.4	6668	14	US-10-311-455-1670
43	124.4	3.4	42533	14	US-10-004-113-43
44	123.6	3.3	2479	15	US-10-108-260A-1966
45	122	3.3	903	14	US-10-029-386-20462

ALIGNMENTS

RESULT 1

US-10-084-817-30
; Sequence 30, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 30
; LENGTH: 3702
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 1312325CB1
US-10-084-817-30

Query Match 100.0%; Score 3702; DB 14; Length 3702;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3702; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTAGGAATGACAGGCATCTCCACAGGAGGTCATCCACCTTGGCTGGGTGCTG 60

Db 1 GCCTAGGAATGACAGGCATCTCCACAGGAGGTCATCCACCTTGGCTGGGTGCTG 60

QY 61 CATTGGCTGCTATTAGAAAAACACAGGACAATGCATACCCGCTCCCGACTGTAA 120

Db 61 CATTGGCTGCTATTAGAAAAACACAGGACAATGCATACCCGCTCCCGACTGTAA 120

QY 121 CATAGGGATATGTGTTCACTTAGCATGSACTTCTGGAGGGCCCAAGAAAGGGCGTCT 180
Db 121 CATAGGGATATGTGTTCACTTAGCATGSACTTCTGGAGGGCCCAAGAAAGGGCGTCT 180
QY 181 GGAAGTTTATGATAGAGAGTGTATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 181 GGAAGTTTATGATAGAGAGTGTATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 CTGGCTGTCTCTGCTTTAAAGAAATCAGTCTCTCTTTCCGACTTAGTCTCTGGGAAGA 300
Db 241 CTGGCTGTCTCTGCTTTAAAGAAATCAGTCTCTCTTTCCGACTTAGTCTCTGGGAAGA 300
QY 301 AGTTTCAGATACAAAGTATCATTTGGAACATTTCAAGATCATCAAAATTCACAGG 360
Db 301 AGTTTCAGATACAAAGTATCATTTGGAACATTTCAAGATCATCAAAATTCACAGG 360
QY 361 GATTGGTGACCAACAGAGGCTCAGACATCTGATTGTCTGACTCTCCAGACATCATCTG 420
Db 361 GATTGGTGACCAACAGAGGCTCAGACATCTGATTGTCTGACTCTCCAGACATCATCTG 420
QY 421 GTCTCCCTGAACTGAATCAACCATGGATGATTTGAGCGTCCGAGGAACCTTAGAAG 480
Db 421 GTCTCCCTGAACTGAATCAACCATGGATGATTTGAGCGTCCGAGGAACCTTAGAAG 480
QY 481 GCAAAAGAGGAGGATGCCACTCGAAGCAGAAAGAAATCGGCTTACCAGAGGAATGACGA 540
Db 481 GCAAAAGAGGAGGATGCCACTCGAAGCAGAAAGAAATCGGCTTACCAGAGGAATGACGA 540
QY 541 TGATGAAGAGAGGAGCGCCCGGAAACGGCGCGCGAGCCGAGAAACGGCTCGCGCA 600
Db 541 TGATGAAGAGAGGAGCGCCCGGAAACGGCGCGCGAGCCGAGAAACGGCTCGCGCA 600
QY 601 GAAGCAGAGGAAGATCTTTGGGACAGTGTACCGACAGAGTGGAGGTGAATGCCAGAA 660
Db 601 GAAGCAGAGGAAGATCTTTGGGACAGTGTACCGACAGAGTGGAGGTGAATGCCAGAA 660
QY 661 CAGTGTCTGACGAGGAGGCAAGCAACCAACCAACCACTCAAGTGGAAAGGGATGA 720
Db 661 CAGTGTCTGACGAGGAGGCAAGCAACCAACCAACCACTCAAGTGGAAAGGGATGA 720
QY 721 TGAGGCGCATCTCTGGAGCGCTGGCTCGGCTGAGGAAGACGCGCAAAACGCTTCA 780
Db 721 TGAGGCGCATCTCTGGAGCGCTGGCTCGGCTGAGGAAGACGCGCAAAACGCTTCA 780
QY 781 GGAGGCTCTGGAGCGGCAAGAGAGTTCGACCCCAACCACTCAAGTGGAAAGGGATGA 840
Db 781 GGAGGCTCTGGAGCGGCAAGAGAGTTCGACCCCAACCACTCAAGTGGAAAGGGATGA 840
QY 841 CCAAGCAGAGAGATGCAAAATGACACAGCAGAAATGAACTACCGAGAGGAAGAA 900
Db 841 CCAAGCAGAGAGATGCAAAATGACACAGCAGAAATGAACTACCGAGAGGAAGAA 900
QY 901 AAGTGAAAGTCCCAAGAAAGATACGAGATAGAGGAAACAGAAACAGTCAACAACTCTA 960
Db 901 AAGTGAAAGTCCCAAGAAAGATACGAGATAGAGGAAACAGAAACAGTCAACAACTCTA 960
QY 961 CCAGAGAAATGATTTGGAGGATGCTGAGAAACCAAGAAAGAGACAGAAAGAGGA 1020
Db 961 CCAGAGAAATGATTTGGAGGATGCTGAGAAACCAAGAAAGAGACAGAAAGAGGA 1020
QY 1021 GGAGGAGAGGAGAAACCAAGCGAGGAGCATTTGGAGAAATCAGGTAGAGGTGATGTT 1080
Db 1021 GGAGGAGAGGAGAAACCAAGCGAGGAGCATTTGGAGAAATCAGGTAGAGGTGATGTT 1080
QY 1081 GGAAGAGAAACCACTGAAGCCAGAGAAACAGTGGTAAATGCTATTAATAATGGGCA 1140
Db 1081 GGAAGAGAAACCACTGAAGCCAGAGAAACAGTGGTAAATGCTATTAATAATGGGCA 1140
QY 1141 GATCAGTTTCAGAGAGCCCTTAAACAGAGGAGAGGAGCAAGGTTTCAGATCAGATTTTC 1200
Db 1141 GATCAGTTTCAGAGAGCCCTTAAACAGAGGAGAGGAGCAAGGTTTCAGATCAGATTTTC 1200
QY 1201 CCATCATGAAAGATGGAAGAGGAACAAAGGAAGAGCTTGAGGCGAGAGGGCAAGGTT 1260

Db 1201 CCATCATGAAAGATGGAAGAGGAAGCAAGGAAAGAGCTGAGGCGAGAGGGCAAGGTT 1260
QY 1261 GGAAGCAGAAAGAGCAAAAGATTTAAAGCCGAGCAACAAAGAGATAGCAGATGAACG 1320
Db 1261 GGAAGCAGAAAGAGCAAAAGATTTAAAGCCGAGCAACAAAGAGATAGCAGATGAACG 1320
QY 1321 AGCAAGATTTGAAGCAGAAAGAGAGAGCTGCCCAAGAAAGAGAGAGAGAGGCGAGA 1380
Db 1321 AGCAAGATTTGAAGCAGAAAGAGAGAGCTGCCCAAGAAAGAGAGAGAGGCGAGA 1380
QY 1381 AGAGAGGGAAGATGAGGAGAGAGAGAAAGGCGCAGCAGAGGAGAGGAGAGATATA 1440
Db 1381 AGAGAGGGAAGATGAGGAGAGAGAGAAAGGCGCAGCAGAGGAGAGGCGAGATATA 1440
QY 1441 GGAAGAGAAAGAAAGGCGCAGAGAGGAGAGAGAGGATATAAGGAGAGAGAGAGAGGCG 1500
Db 1441 GGAAGAGAAAGAAAGGCGCAGAGAGGAGAGAGAGGATATAAGGAGAGAGAGAGAGGCG 1500
QY 1501 AGCAGAGGAGAGGAGAGGATATAAGAGAGAGAGAGAGAGGCGCAGAGAGAGGCGAAAG 1560
Db 1501 AGCAGAGGAGAGGAGAGGATATAAGAGAGAGAGAGAGAGGCGCAGAGAGAGGCGAAAG 1560
QY 1561 GGCCAGGCGCAGAGAGAGAGAGAGGCTTAAGGTAGAGAGAGAGAGAGAGAGAGCT 1620
Db 1561 GGCCAGGCGCAGAGAGAGAGAGAGGCTTAAGGTAGAGAGAGAGAGAGAGAGAGCT 1620
QY 1621 AGAAGAGAAAGAAACATCGCATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1621 AGAAGAGAAAGAAACATCGCATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1681 AATAGAGGGGAAATGGGTAAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGCTGT 1740
Db 1681 AATAGAGGGGAAATGGGTAAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGCTGT 1740
QY 1741 CCTTAAAGCTCCA 1800
Db 1741 CCTTAAAGCTCCA 1800
QY 1801 AGAAGCAAGCTTACCTTCAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Db 1801 AGAAGCAAGCTTACCTTCAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
QY 1861 AGAAGCAAGTAA 1920
Db 1861 AGAAGCAAGTAA 1920
QY 1921 GTCCAGAGATGAGAGATTCATGACCCCAAACTTAAACATCTGAGAGATCTTTAGCCG 1980
Db 1921 GTCCAGAGATGAGAGATTCATGACCCCAAACTTAAACATCTGAGAGATCTTTAGCCG 1980
QY 1981 CCCTGGAGGAGGCGAGCTGGACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040
Db 1981 CCCTGGAGGAGGCGAGCTGGACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040
QY 2041 CGGCAAGAGGCTGGAGAGAGCTTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 2041 CGGCAAGAGGCTGGAGAGAGCTTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
QY 2101 GAAAGCTCAAAAG 2160
Db 2101 GAAAGCTCAAAAG 2160
QY 2161 GGAGGAG 2220
Db 2161 GGAGGAG 2220
QY 2221 TCGAAAGCTCAG 2280
Db 2221 TCGAAAGCTCAG 2280
QY 2281 AGCAGAGCTGCTGAG 2340

2281 AGCAGAACTGCTGAGAAAGCGCAGAGATGCCAGAGATGGCTGTGATGACAGAGAA 2340
2341 ACCATTCAAGTGTTTCACTCTCTAAAGGTTTCATCTCTCAAGATAGAGAGAGCAGCAATT 2400
2341 ACCATTCAAGTGTTTCACTCTCTAAAGGTTTCATCTCTCAAGATAGAGAGAGCAGCAATT 2400
2401 TTTGAAATAGTCTGTCAGAAAGCGTGGTGTCAATCGACCAATCAAGCAGCAGATAGT 2460
2401 TTTGAAATAGTCTGTCAGAAAGCGTGGTGTCAATCGACCAATCAAGCAGCAGATAGT 2460
2461 CTCCAAGATTGACAGCAGACTGGAGCAGTATACAGTGCATTTGAGGGAACAAAAGCGC 2520
2461 CTCCAAGATTGACAGCAGACTGGAGCAGTATACAGTGCATTTGAGGGAACAAAAGCGC 2520
2521 AAAACCTCAAAAGCGCGACCTCGATCTCTCTGTTCTCTGCTGAGGTGTACCAACAT 2580
2521 AAAACCTCAAAAGCGCGACCTCGATCTCTCTGTTCTCTGCTGAGGTGTACCAACAT 2580
2581 CAAGAGTATGTCGGAGAAAGGAATGTGTTTTCATCCCCCACTGCAGCAGCAGCAACAAA 2640
2581 CAAGAGTATGTCGGAGAAAGGAATGTGTTTTCATCCCCCACTGCAGCAGCAGCAACAAA 2640
2641 TAAGAAACTGCTGCTTGAAGTGAAGTGGGTTTCTAGCCGATCAATGAATGGCTAACTAA 2700
2641 TAAGAAACTGCTGCTTGAAGTGAAGTGGGTTTCTAGCCGATCAATGAATGGCTAACTAA 2700
2701 AACCCAGATGGAACAAGTCACTCTCTCCCAACCTTCTGACTTGAGACAGGAGACGT 2760
2701 AACCCAGATGGAACAAGTCACTCTCTCCCAACCTTCTGACTTGAGACAGGAGACGT 2760
2761 ATCCAGAGCGGAACCTCTGGGAAAGCAATCTGTGATGAAGTCACTTCCCCCACTAA 2820
2761 ATCCAGAGCGGAACCTCTGGGAAAGCAATCTGTGATGAAGTCACTTCCCCCACTAA 2820
2821 GGTGTTGAGACAGTTCAGAAAGAACCCAGCTCAAGCAGCAGCAGCAGCTCAGTTGAGA 2880
2821 GGTGTTGAGACAGTTCAGAAAGAACCCAGCTCAAGCAGCAGCAGCAGCTCAGTTGAGA 2880
2881 GGGCTAAATCGCTCTGTTTGTATTTATGTTGATTTACTAAATGGGTTCAATTCATTTT 2940
2881 GGGCTAAATCGCTCTGTTTGTATTTATGTTGATTTACTAAATGGGTTCAATTCATTTT 2940
2941 ATTTTCAATATCCAGTAAACCCATGATATATATCACTATATTTAATATCAAGTCTA 3000
2941 ATTTTCAATATCCAGTAAACCCATGATATATATCACTATATTTAATATCAAGTCTA 3000
3001 GAGATGTTTCATGTAAAGTACTGCTTTCACAGGAGCCTGTTTCTAAAGAAACCCCATG 3060
3001 GAGATGTTTCATGTAAAGTACTGCTTTCACAGGAGCCTGTTTCTAAAGAAACCCCATG 3060
3061 CTGTGAAATAGAGACTTTTCTACTGATCATATAAATCTCTGATCTGAGCAGTATACCAA 3120
3061 CTGTGAAATAGAGACTTTTCTACTGATCATATAAATCTCTGATCTGAGCAGTATACCAA 3120
3121 CCACATCTGAAGTCAACAGAAATCCAAATTTAAATGCTCGGGAATGTGTCAGTAT 3180
3121 CCACATCTGAAGTCAACAGAAATCCAAATTTAAATGCTCGGGAATGTGTCAGTAT 3180
3181 CTAGAAAAATGAACCGTAGTTTGTGTTTAAATACAGAAATCATGTTGTTTCTGCAAC 3240
3181 CTAGAAAAATGAACCGTAGTTTGTGTTTAAATACAGAAATCATGTTGTTTCTGCAAC 3240
3241 TTTATATTAAGCATGGAGAAATTAATCTAGTAGGCAATTTGACACTTTTGAAGTA 3300
3241 TTTATATTAAGCATGGAGAAATTAATCTAGTAGGCAATTTGACACTTTTGAAGTA 3300
3301 ACCCATTTTCAAGTTGAAATCTGCGATAATGTTGCTTTAAAAAAGAAAAAGATGC 3360
3301 ACCCATTTTCAAGTTGAAATCTGCGATAATGTTGCTTTAAAAAAGAAAAAGATGC 3360
3361 GGGCAGATTAATCTTTGTTGGGTAAATTTACTTGACCTGGCCGGTTACAGTCTGTG 3420
3361 GGGCAGATTAATCTTTGTTGGGTAAATTTACTTGACCTGGCCGGTTACAGTCTGTG 3420

3421 CTGGGAAACTTGGCGTCCCAATTAATCGCATGTCCGAATCCCTTCGACGCTGGTATGCGAA 3480
3421 CTGGGAAACTTGGCGTCCCAATTAATCGCATGTCCGAATCCCTTCGACGCTGGTATGCGAA 3480
3481 GGGCAGCATGGCTCCAAATGCGGTGATGGATGGGGCTGTGCTTACGCGGTTTGGGT 3540
3481 GGGCAGCATGGCTCCAAATGCGGTGATGGATGGGGCTGTGCTTACGCGGTTTGGGT 3540
3541 TCCGACCGAAATTCAGCACCGGGCTTGTCTCTTTTGGGGGGGGCCCTAAAGGGCTTG 3600
3541 TCCGACCGAAATTCAGCACCGGGCTTGTCTCTTTTGGGGGGGGCCCTAAAGGGCTTG 3600
3601 GGGTTTTTCCGCCGAAATTTGGCTTGTGCCGAAATCCCTTTGGGGTTACGCTCAGACAC 3660
3601 GGGTTTTTCCGCCGAAATTTGGCTTGTGCCGAAATCCCTTTGGGGTTACGCTCAGACAC 3660
3661 GTTTTGTCCGGGAGAGGAAGCAAAAGGGGGGTTACCGAGAAC 3702
3661 GTTTTGTCCGGGAGAGGAAGCAAAAGGGGGGTTACCGAGAAC 3702

RESULT 2

US-10-062-674-1815
; Sequence 1815, Application US/10062674
; Publication No. US2004000559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 09/625,102
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 2217
; SOFTWARE: PEEL Program
; SEQ ID NO 1815
; LENGTH: 3500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US2004000559A1 256147.56
; NAME/KEY: unsure
; LOCATION: (1) ... (3500)
; OTHER INFORMATION: a, t, c, g, or other
US-10-062-674-1815

Query Match 86.5%; Score 3201.2; DB 15; Length 3500;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 3206; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

3421 GGGATATGTGTTTCACTTAGCATGACTTCTGGAGGGGCCAAGGAAGGGCGGTCTGGAGT 185
1 GGGATATGTGTTTCACTTAGCATGACTTCTGGAGGGGCCAAGGAAGGGCGGTCTGGAGT 60
186 TTTATTGAATAGCAGTGTGTTATTCGGCTGCTGCCCTGCCCTGCTCTCTCTGCTG 245
61 TTTATTGAATAGCAGTGTGTTATTCGGCTGCTGCCCTGCCCTGCTCTCTCTGCTG 120
246 TGTGCTCTCTGCTTTAAAGAAATCAGTCTCTTTCCGACTTAGTCTCGGGAAGAGTTT 305
121 TGTGCTCTCTGCTTTAAAGAAATCAGTCTCTTTCCGACTTAGTCTCGGGAAGAGTTT 180
306 CAGACTACAAGTATCATTTGGAACATTTCAAGATCATCAATCAATCAATCCACAGGATG 365
181 CAGACTACAAGTATCATTTGGAACATTTCAAGATCATCAATCAATCAATCCACAGGATG 240
366 GTGACCAACCAAGGCTCAGACATCTGATTTGCTGACCTGTCCAGACATCATCTGCTCTC 425
241 GTGACCAACCAAGGCTCAGACATCTGATTTGCTGACCTGTCCAGACATCATCTGCTCTC 300

426 QY CCTGAACCTGAAATCACACCATGGATGATTTTGTAGCGTTCGAGAGAACTTGAAGGCCAAA 485
301 Db CCTGAACCTGAAATCACACCATGGATGATTTTGTAGCGTTCGAGAGAACTTGAAGGCCAAA 360
486 QY AGAGGGAGGAGATCGGACTCGAGCAGAGAAAGAAATCGCTTACAGAGGAATGACGATG 545
361 Db AGAGGGAGGAGATCGGACTCGAGCAGAGAAAGAAATCGCTTACAGAGGAATGACGATG 420
546 QY AAGAGGAGGACGCCGGGAAACGGCCGCGCGAGCGCGAGCAAGAAACGGCTCGCGCAAGAAC 605
421 Db AAGAGGAGGACGCCGGGAAACGGCCGCGCGAGCGCGAGCAAGAAACGGCTCGCGCAAGAAC 480
506 QY AGAGGAGAGATCTTGTGGACAGGTGACCGCAGCGAGTGCAGTGCAGTGCAGCAACAGTG 665
481 Db AGAGGAGAGATCTTGTGGACAGGTGACCGCAGCGAGTGCAGTGCAGTGCAGCAACAGTG 540
666 QY TGCCTGACGAGAGGAGCCCAAGCAACACCAACCAACCACTCAAGTGGAGGGGATGATGAGG 725
541 Db TGCCTGACGAGAGGAGCCCAAGCAACACCAACCAACCACTCAAGTGGAGGGGATGATGAGG 600
726 QY CCGCATTCCTGAGAGCGCTTGGCTCGGCTGAGGAAGACCGCAAAACGGCTTCAGGAGG 785
601 Db CCGCATTCCTGAGAGCGCTTGGCTCGGCTGAGGAAGACCGCAAAACGGCTTCAGGAGG 660
786 QY CTCTGGAGCGGAGAGGAGTTCGACCCCAACCAATACAGATGCAAGTCTGTCTGCCCAA 845
661 Db CTCTGGAGCGGAGAGGAGTTCGACCCCAACCAATACAGATGCAAGTCTGTCTGCCCAA 720
846 QY GCAGAAGAAATGCAAAATGACACAGCAGAAATTAAGAACTACCGAGAGGAAGAAAGATG 905
721 Db GCAGAAGAAATGCAAAATGACACAGCAGAAATTAAGAACTACCGAGAGGAAGAAAGATG 780
906 QY AAAGTCGCCAAGAAAGATACGAGATAGAGAAACAGAAACAGTCAACCAAGTCTTACCAGA 965
781 Db AAAGTCGCCAAGAAAGATACGAGATAGAGAAACAGAAACAGTCAACCAAGTCTTACCAGA 840
966 QY AGAATGATTTGGAGGATGCTGAAGAAACAAAGAAAGAGCAAGAAAGAGAGGAGG 1025
841 Db AGAATGATTTGGAGGATGCTGAAGAAACAAAGAAAGAGCAAGAAAGAGAGGAGGAGG 900
1026 QY AAGAGGAGAGCCAAAGCCAGGAGGAGTTCGAGAAATACAGGTAGAGGTGATGCTGGAAG 1085
901 Db AAGAGGAGAGCCAAAGCCAGGAGGAGTTCGAGAAATACAGGTAGAGGTGATGCTGGAAG 960
1086 QY AGAAACCAACTGAAGCCAGGAGGAAACAGTGTATGTCTATTAATAATGGGCAGATCA 1145
961 Db AGAAACCAACTGAAGCCAGGAGGAAACAGTGTATGTCTATTAATAATGGGCAGATCA 1020
1146 QY GTTCAGAGAGCTTAACAGAGGAGGAGGAGGAAACAGGTTCAGATGAGATTTCCCATC 1205
1021 Db GTTCAGAGAGCTTAACAGAGGAGGAGGAGGAAACAGGTTCAGATGAGATTTCCCATC 1080
1206 QY ATGAAAGATGGAAGAGGAGGAGCAAGGAAAGAGCTGAGGCAGAGAGGCGCAAGTTGGAAG 1265
1081 Db ATGAAAGATGGAAGAGGAGGAGCAAGGAAAGAGCTGAGGCAGAGAGGCGCAAGTTGGAAG 1140
1266 QY CAGAGAAAGAGAAAGAAATTAAGCCGAGCAAGCAAAAGAGATAGCAGATGAACGAGCAA 1325
1141 Db CAGAGAAAGAGAAAGAAATTAAGCCGAGCAAGCAAAAGAGATAGCAGATGAACGAGCAA 1200
1326 QY GAATTTGAAGCAGAAAGAAAGAGCTGCCCAAGAAAGAGAGAGAGAGGAGGAGAGAGA 1385
1201 Db GAATTTGAAGCAGAAAGAAAGAGCTGCCCAAGAAAGAGAGAGAGAGGAGGAGAGAGA 1260
1386 QY GGGAAAGATGAGGAGGAGGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1445
1261 Db GGGAAAGATGAGGAGGAGGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
1446 QY AAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1505
1321 Db AAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
1506 QY AGG 1565

1381 Db AGGAGGCGAGAGATATAAGAGGAGAGAAAGGGCGAGAGGAGGAGGAGGAGGAGGAGG 1440
1566 QY GGGCAGAGGAGGAGAGAGAGGCTTAAGGTAGAGGAGCAGAGAAACGTAACAAGAGCTAGAAG 1625
1441 Db GGGCAGAGGAGGAGAGAGAGGCTTAAGGTAGAGGAGCAGAGAAACGTAACAAGAGCTAGAAG 1500
1626 QY AGAAAAACATGCCATGCCATGCAAGAGCAAAAGATAAAGGGGAAAGGTAGAACAGAAAAATAG 1685
1501 Db AGAAAAACATGCCATGCCATGCAAGAGCAAAAGATAAAGGGGAAAGGTAGAACAGAAAAATAG 1560
1686 QY AAGGGAATGGGTAAATGAAAGAAAGCAACAGAGATATAAATTCACAGACAGTGTCTCTAA 1745
1561 Db AAGGGAATGGGTAAATGAAAGAAAGCAACAGAGATATAAATTCACAGACAGTGTCTCTAA 1620
1746 QY AGAAACAGGAGAGAGAGAGGAGGAACTAAAGTCCAGCTTAAAGAGAAAGAGTCCCAAGAG 1805
1621 Db AGAAACAGGAGAGAGAGAGGAGGAACTAAAGTCCAGCTTAAAGAGAAAGAGTCCCAAGAG 1680
1806 QY ACAAGCCTTACCTTCAAAAAAGAGAGATCAAAAGATGAAAGATTAATAAGGACAAAGAAC 1865
1681 Db ACAAGCCTTACCTTCAAAAAAGAGAGATCAAAAGATGAAAGATTAATAAGGACAAAGAAC 1740
1866 QY CCAAGAGAAAGATTAAGAGCTTCATGATCGAAAGAGGATTTACAGAAATTAAGTCGC 1925
1741 Db CCAAGAGAAAGATTAAGAGCTTCATGATCGAAAGAGGATTTACAGAAATTAAGTCGC 1800
1926 QY AGAATGGAGATTCATGATCACCAAACTTAAACATACATCAGATACTTTTCAGCGCCCTG 1985
1801 Db AGAATGGAGATTCATGATCACCAAACTTAAACATACATCAGATACTTTTCAGCGCCCTG 1860
1986 QY GAGGAGGCGCAGCGTGGACACCAAGAGGCTTGAGGCGCCCGCCCGAGGTGGAAGCCGGCA 2045
1861 Db GAGGAGGCGCAGCGTGGACACCAAGAGGCTTGAGGCGCCCGCCCGAGGTGGAAGCCGGCA 1920
2046 QY AAAGCTCGAGAGCTTCGTCTCGTCGCGGAGACCGAGAGCGAGAGTTCAGAGAGC 2105
1921 Db AAAGCTCGAGAGCTTCGTCTCGTCGCGGAGACCGAGAGCGAGAGTTCAGAGAGC 1980
2106 QY TCAACAGAGCAGCAGGAGGCGGCTTTGGAGCTGAGAGAACTCAAGAAAAAGAGGAGG 2165
1981 Db TCAACAGAGCAGCAGGAGGCGGCTTTGGAGCTGAGAGAACTCAAGAAAAAGAGGAGG 2040
2166 QY AGAGAGGAGAGCTCTGGAGGAGGAGAGCAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2225
2041 Db AGAGAGGAGAGCTCTGGAGGAGGAGAGAGCAGCAGAGGAGGAGGAGGAGGAGGAGGAGG 2100
2226 QY AACTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2285
2101 Db AACTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
2286 QY AAGCTCTGAGAAACGCCAGAGAGTCCAGAGAGATGCTTGTGATGAGCAAGAAACCAT 2345
2161 Db AAGCTCTGAGAAACGCCAGAGAGTCCAGAGAGATGCTTGTGATGAGCAAGAAACCAT 2220
2346 QY TCAAGTGTCTTCACTCTCTAAAGGTTTCATCTCTCAAGATAGAAGAGCGAGCAATTTTGA 2405
2221 Db TCAAGTGTCTTCACTCTCTAAAGGTTTCATCTCTCAAGATAGAAGAGCGAGCAATTTTGA 2280
2406 QY ATAAGTCTGTGAGAAACAGCAGTGTGTCAATTCGACCCATCAGCAGCATATGCTCCA 2465
2281 Db ATAAGTCTGTGAGAAACAGCAGTGTGTCAATTCGACCCATCAGCAGCATATGCTCCA 2340
2466 QY AGATTGACAGCAGACTGGAGCAGTATACCAAGTGCATTTGAGGGAACAAAAAGCCAAAAAC 2525
2341 Db AGATTGACAGCAGACTGGAGCAGTATACCAAGTGCATTTGAGGGAACAAAAAGCCAAAAAC 2400
2526 QY CTACAAAGCCGCGAGCTCGATCTTCTGTTCTGCTGAGGTGTAGCGCACTCAGA 2585
2401 Db CTACAAAGCCGCGAGCTCGATCTTCTGTTCTGCTGAGGTGTAGCGCACTCAGA 2460
2586 QY GTATGTGGGAGAAAGGGAATGTGTTTTCATCTCCCTGAGGAGGAGGAGGAGGAGGAGG 2645

QY 1837 AGATGAAGAATTAAAGGACAAAGAACCCAAAGAGAGTTAAGAGCTTCATGGATCG 1896
 Db 1621 AGATGAAGAATTAAAGGACAAAGAACCCAAAGAGAGTTAAGAGCTTCATGGATCG 1680
 QY 1897 AAAGAGGAGTTTACAGAGTTTAAAGTCGAGAGTTTAAAGTTTAAAGTTTAAAGTTTAA 1956
 Db 1681 AAAGAGGAGTTTACAGAGTTTAAAGTCGAGAGTTTAAAGTTTAAAGTTTAAAGTTTAA 1740
 QY 1957 ACATAGTGAATATCTTTCAGCGCCCTTGAGGAGGCGGAGGCTGAGACCAAGAGGCG 2016
 Db 1741 ACATAGTGAATATCTTTCAGCGCCCTTGAGGAGGCGGAGGCTGAGACCAAGAGGCG 1800
 QY 2017 TGAGGCGCCCTTGAGGAGGCGGAGGCTGAGAGGCTGAGAGGCTTCTGCTGCTGCGG 2076
 Db 1801 TGAGGCGCCCTTGAGGAGGCGGAGGCTGAGAGGCTTCTGCTGCTGCGG 1860
 QY 2077 GGAGACCGAGAGGAGGAGTTTCAAGAGGCTTCAACAGAGGAGGAGGAGGCTTCTGGA 2136
 Db 1861 GGAGACCGAGAGGAGGAGTTTCAAGAGGCTTCAACAGAGGAGGAGGAGGCTTCTGGA 1920
 QY 2137 GCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2196
 Db 1921 GCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1980
 QY 2197 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2256
 Db 1981 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2040
 QY 2257 AAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2316
 Db 2041 AAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
 QY 2317 AGATGGCTTCTGAGTGAACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2376
 Db 2101 AGATGGCTTCTGAGTGAACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
 QY 2377 CAAGATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2436
 Db 2161 CAAGATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
 QY 2437 ATCGACCCATCAAGCAGCAATAGTCTCAAGATTTGAGAGGAGGAGGAGGAGGAGGAGG 2496
 Db 2221 ATCGACCCATCAAGCAGCAATAGTCTCAAGATTTGAGAGGAGGAGGAGGAGGAGGAGG 2280
 QY 2497 TGCAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2556
 Db 2281 TGCAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
 QY 2557 TCCTGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2616
 Db 2341 TCCTGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2400
 QY 2617 CCCCACTGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2676
 Db 2401 CCCCACTGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2460
 QY 2677 CCGCATCAATGAATGGCTTAACTTAAACCCAGATGGAAACCAAGTCACTCTCCCAACC 2736
 Db 2461 CCGCATCAATGAATGGCTTAACTTAAACCCAGATGGAAACCAAGTCACTCTCCCAACC 2520
 QY 2737 TTCTGACTTGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2796
 Db 2521 TTCTGACTTGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2580
 QY 2797 GGATAAGGTCATCTCCCACTAAGTTTGAACAGTTTCAAGAGGAGGAGGAGGAGGAGGAGG 2856
 Db 2581 GGATAAGGTCATCTCCCACTAAGTTTGAACAGTTTCAAGAGGAGGAGGAGGAGGAGGAGG 2640
 QY 2857 ACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2916
 Db 2641 ACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2700

QY 2917 ACTAAATGGGTTTCAATATCTTTTCAATATCCAGTAAACCCATGTATATATC 2976
 Db 2701 ACTAAATGGGTTTCAATATCTTTTCAATATCCAGTAAACCCATGTATATATC 2760
 QY 2977 ACTAATTTAATAATCACAGTCTAGAGATGTTTCATGTTAAAGTACTGCTTTCACAGG 3036
 Db 2761 ACTAATTTAATAATCACAGTCTAGAGATGTTTCATGTTAAAGTACTGCTTTCACAGG 2820
 QY 3037 AGCCTGTTTCTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATATAAC 3096
 Db 2821 ATCCCTGTTTCTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATATAAC 2880
 QY 3097 TCTGTATCTGAGCAGTGTATCCAAACCATCTCTGAAGTCAACAGAGATCCAAAGTTTAAA 3156
 Db 2881 TCTGTATCTGAGCAGTGTATCCAAACCATCTCTGAAGTCAACAGAGATCCAAAGTTTAAA 2940
 QY 3157 TTGCTCTGCGGAATGCTGAGTATCTAGAAATGAACCGTAGTTTGTGTTTTTAAAT 3216
 Db 2941 TTG-CTGCGGAATGCTGAGTATCTAGAAATGAACCGTAGTTTGTGTTTTTAAAT 2999
 QY 3217 ACAGAGTCAATGTTTCTGCACTTTATTAATTAAGCATCGAAGATTTCTTAGTAGG 3276
 Db 3000 ACAGAGTCAATGTTTCTGCACTTTATTAATTAAGCATCGAAGATTTCTTAGTAGG 3059
 QY 3277 CAATTGTAACATTTTGAAGTAAACCATTTTGAAGTAAACCATTTTGAAGTAAACCATTT 3336
 Db 3060 CAATTGTAACATTTTGAAGTAAACCATTTTGAAGTAAACCATTTTGAAGTAAACCATTT 3119
 QY 3337 CTTTAAAAAAGAAAAAGATCGGCGACAGTTATTCCTTT 3378
 Db 3120 TCTTTAAAAAAGAAAAAGATGTTACTGTTAAGTATTACTTTT 3161

RESULT 5

US-10-062-674-1817/C
 ; Sequence 1817, Application US/10062674
 ; Publication No. US20040005559A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
 ; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
 ; FILE REFERENCE: PA-0026-1 CIP
 ; CURRENT APPLICATION NUMBER: US/10/062,674
 ; PRIOR FILING DATE: 2002-01-30
 ; PRIOR APPLICATION NUMBER: US 09/625,102
 ; PRIOR FILING DATE: 2000-07-24
 ; NUMBER OF SEQ ID NOS: 2217
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 1817
 ; LENGTH: 4566
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: incyte ID No. US20040005559A1 256147.58C
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1) ... (4566)
 ; OTHER INFORMATION: a, t, c, g, or other
 ; US-10-062-674-1817

Query Match 63.6%; Score 2353; DB 15; Length 4566;
 Best Local Similarity 92.0%; Pred. No. 0;
 Matches 2616; Conservative 0; Mismatches 11; Indels 217; Gaps 4;
 QY 517 AATCGCTTACCAGAGGAGTACGATGATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 576
 Db 4474 AATCGCTTACCAGAGGAGTACGATGATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4416
 QY 577 AGCCCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 636
 Db 4415 AGCCCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4356
 QY 637 CCAGTGGAGGTGAATGCCAGAAACAGTGTGCTGTGACGAGGAGGAGGAGGAGGAGGAGGAGG 696

||||| 4355 CCAGGTGGAGGTGAATGCCAGAAACAGTGTGCTGACAGAGGAGGCGCAACACACACAC 4296
 ||||| 697 AACACCT - CAACTGGAAGGGGATGATGAGGCGCGATTCCTGAGAGGCCCTGGGTCTGGCGGTG 755
 ||||| 4295 AACACCTGCAAGTGGAGGGGATGATGAGGCGCGATTCCTGAGAGGCCCTGGGTCTGGCGGTG 4236
 ||||| 756 AGGAAAGACGCCAAACACCGCTTCAGGAGGCTCTGGAGCGGCGAGAGGATTCGACCCAA 815
 ||||| 4235 AGGAAAGACGCCAAACACCGCTTCAGGAGGCTCTGGAGCGGCGAGAGGATTCGACCCAA 4176
 ||||| 816 CAATAACAGATGCAAGTCTGTGCTCCCAAGCAGAGAAATGCAAAATGACACAGCAGAAA 875
 ||||| 4175 CAATAACAGATGCAAGTCTGTGCTCCCAAGCAGAGAAATGCAAAATGACACAGCAGAAA 4116
 ||||| 876 ATGAAACTACCGAGAGGAAGAAAAGTGAAGTGCCTCCCAAGAGATACGAGATAGAG 935
 ||||| 4115 ATGAAACTACCGAGAGGAAGAAAAGTGAAGTGCCTCCCAAGAGATACGAGATAGAG 4056
 ||||| 936 AACACAGAAAACAGTACCAGTCTCTACCAAGAAATGATTTGGAGGGATGCTGAAGAAAACA 995
 ||||| 4055 AACACAGAAAACAGTACCAGTCTCTACCAAGAAATGATTTGGAGGGATGCTGAAGAAAACA 3996
 ||||| 996 AGAAGAGAGACAAGAAAAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1055
 ||||| 3995 AGAAGAGAGACAAGAAAAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3936
 ||||| 1056 GAGAAAATCAGGTAGAGGTGATGGTGGAGAGAAAACAACTGAAAGCCAGGAGAGAAACAG 1115
 ||||| 3935 GAGAAAATCAGGTAGAGGTGATGGTGGAGAGAAAACAACTGAAAGCCAGGAGAGAAACAG 3876
 ||||| 1116 TGGTAATGTCTATTAATAAATGGGCGAGATCAGTTTCAAGAGAGCTTAACAGAGAGAGAGAG 1175
 ||||| 3875 TGGTAATGTCTATTAATAAATGGGCGAGATCAGTTTCAAGAGAGCTTAACAGAGAGAGAGAG 3816
 ||||| 1176 GGGAAACAGGTTTCAATGAGATTTCCCATCATGAAAAGATGGAAGAGGAGAGACAAGGAAA 1235
 ||||| 3815 GGGAAACAGGTTTCAATGAGATTTCCCATCATGAAAAGATGGAAGAGGAGAGACAAGGAAA 3756
 ||||| 1236 GAGCTGAGCGAGAGAGGCGCAAGGTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1295
 ||||| 3755 GAGCTGAGCGAGAGAGGCGCAAGGTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3696
 ||||| 1296 AAGACAAAAGATACAGATGAACAGAGCAAGATTTCAAGCAGAGAGAGAGAGAGAGAGAGAGAG 1355
 ||||| 3695 AAGACAAAAGATACAGATGAACAGAGCAAGATTTCAAGCAGAGAGAGAGAGAGAGAGAGAGAG 3636
 ||||| 1356 AAGAAAGAGAAAGGAG 1415
 ||||| 3635 AAGAAAGAGAAAGGAG 3576
 ||||| 1416 CAGCAGAGAGAGAGAGAGAGATTAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1475
 ||||| 3575 CAGCAGAGAGAGAGAGAGATTAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3516
 ||||| 1476 GGATTAAG 1535
 ||||| 3515 GGATTAAG 3456
 ||||| 1536 AAAGGGCGCAG 1595
 ||||| 3455 AAAGGGCGCAG 3396
 ||||| 1596 AAGACAGAGAAACGTAACAAGCAGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1655
 ||||| 3395 AAGACAGAGAAACGTAACAAGCAGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3336
 ||||| 1656 TAAAGGGGAG 1715
 ||||| 3335 TAAAGGGGAG 3276
 ||||| 1716 AAGAGATATAACTTCAGACAGCTGTCTCTAAAGAAAACAGGAGAGAGAGAGAGAGAGAGAGAG 1775

Db 3275 AAGNAGATAAACTTCAGACAGCTGCTCTAAAGAAACAGGAGAGAGAGAGAGAGAGAGAGAG 3216
 Qy 1776 TCGAAGCTAAAGAGAGAAAGCTCCAAAGAGACAGACCTACCTTCAAAAAAGAGAGATCA 1835
 Db 3215 TCGAAGCTAAAGAGAGAAAGCTCCAAAGAGACAGACCTACCTTCAAAAAAGAGAGATCA 3156
 Qy 1836 AAGATGAAAAGATTAAAG 1895
 Db 3155 AAGATGAAAAGATTAAAG 3096
 Qy 1896 GAAAGAGAGGATTTACAGAGATTAAAGTCGAGAAATGGAGAAATTCATGACCCACAAACTTA 1955
 Db 3095 GAAAGAGAGGATTTACAGAGATTAAAGTCGAGAAATGGAGAAATTCATGACCCACAAACTTA 3036
 Qy 1956 AACATCTGAGAAATATTTTACAGCCGCTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2015
 Db 3035 AACATCTGAGAAATATTTTACAGCCGCTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2976
 Qy 2016 CTGAGGGGCGCCCGCAGGTGGAAGCCGCAAAAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2075
 Db 2975 CTGAGGGGCGCCCGCAGGTGGAAGCCGCAAAAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2916
 Qy 2076 GGGAGACCGAG 2135
 Db 2915 GGGAGACCGAG 2856
 Qy 2136 AGCTGGAGGAACTCAGAAAAG 2195
 Db 2855 AGCTGGAGGAACTCAGAAAAG 2796
 Qy 2196 AGAGGAG 2255
 Db 2795 AGAGGAG 2736
 Qy 2256 TAAAGAGAGAGATTGAAAAGGCGAG 2315
 Db 2735 TAAAGAGAGAGATTGAAAAGGCGAG 2676
 Qy 2316 AAGATGGCTTGTCAATGACAG 2375
 Db 2675 AAGATGGCTTGTCAATGACAG 2616
 Qy 2376 TCAAGATAG 2435
 Db 2615 TCAAGATAG 2556
 Qy 2436 AATCGACCCATCAAGCAGCAATAGTCTCCAAAGATTGACAGCAGACTGGAGCAGTATACCA 2495
 Db 2555 AATCGACCCATCAAGCAGCAATAGTCTCCAAAGATTGACAGCAGACTGGAGCAGTATACCA 2496
 Qy 2496 GTGCAATTTGAGGGAACAAAAGCGGCAAAACCTTCAAAAGCCGCGCAGCTCGAGTCTTCTTG 2555
 Db 2495 GTGCAATTTGAGGGAACAAAAGCGGCAAAACCTTCAAAAGCCGCGCAGCTCGAGTCTTCTTG 2437
 Qy 2556 TTCTCTGTGAGGTGTGACCAACATCAAGAGTATGTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2615
 Db 2436 TTCTCTGTGAGGTGTGACCAACATCAAGAGTATGTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2377
 Qy 2616 CCCCACTGTGAGCAGGCAACCAAAATTAAGAGAAATCTGCTGCTTGAAGGTAGGGGTTCCTA 2675
 Db 2376 CCCCACTGTGAGCAGGCAACCAAAATTAAGAGAAATCTGCTGCTTGAAGGTAGGGGTTCCTA 2317
 Qy 2676 GCGCATCAATGAATGGCTTAACTTAAACCCAGATGGAACCAAGTCAACCTGCTCCCAACAC 2735
 Db 2316 GCGCATCAATGAATGGCTTAACTTAAACCCAGATGGAACCAAGTCAACCTGCTCCCAACAC 2257
 Qy 2736 TTCTGTGAGCAGGCAACCAAAATTAAGAGAAATCTGCTGCTTGAAGGTAGGGGTTCCTA 2795
 Db 2256 TTCTGTGAGCAGGCAACCAAAATTAAGAGAAATCTGCTGCTTGAAGGTAGGGGTTCCTA 2197
 Qy 2796 TCGATAAGGTCTACTTCCCGCAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2855
 Db 2196 TCGATAAGGTCTACTTCCCGCAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2173

```

QY 2856 GAGCAGGACGAGCTCAGTTGTATAGAGGCTAAATCGCTCTGTTTGTATTTATGTTGANT 2915
Db 2172 ----- 2173
QY 2916 TACTAAATGGGTTTCAATATCTTTTATTTTCAATATCCCAAGTAAACCCATGATATATTAT 2975
Db 2172 ----- 2173
QY 2976 CACTATATTAATTAATACACAGTCTAGAGATGTTTCATGGTAAAGTACTGCGCTTGCACAG 3035
Db 2172 -----AG 2171
QY 3036 GAGCCTGTTTCTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGATCATATAA 3095
Db 2170 GAGCCTGTTTCTAAAGAAACCCATGTTGTGAATAGAGACTTTTCTACTGATCATATAA 2111
QY 3096 CTCTGATCTGACAGTGATACCAACACATCTCTGAAGTCAACAGAGATCCAAAGTTTAAA 3155
Db 2110 CTCTGATCTGACAGTGATACCAACACATCTCTGAAGTCAACAGAGATCCAAAGTTTAAA 2051
QY 3156 ATTGCGCTGCGGAATGCTGAGTATCTAGAAAATGAACCGTAGTTTTTGTGTTTTTAAA 3215
Db 2050 ATTGCGCTGCGGAATGCTGAGTATCTAGAAAATGAACCGTAGTTTTTGTGTTTTTAAA 1991
QY 3216 TACAGAGTCTATGTTGTTTCTGCACTTTATATAAAGCATGGAAGAAATATCTTTAGTAG 3275
Db 1990 TACAGAGTCTATGTTGTTTCTGCACTTTATATAAAGCATGGAAGAAATATCTTTAGTAG 1931
QY 3276 GCAATGTACACTTTTGAAGTAAACCATTTTCAGATTTGAAATACGCGATATGTT 3335
Db 1930 GCAATGTACACTTTTGAAGTAAACCATTTTCAGATTTGAAATACGCGATATGTT 1871
QY 3336 GCTTTTAAAAAAGAAAAAGATG 3359
Db 1870 GTCTTTAAAAAAGAAAAAGATG 1847

RESULT 6
US-09-974-298-36
; Sequence 36, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen Hui-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 4547
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 022404.25
US-09-974-298-36

Query Match 48.4%; Score 1793.6; DB 9; Length 4547;
Best Local Similarity 76.9%; Pred. No. 0;
Matches 2585; Conservative 0; Mismatches 9; Indels 766; Gaps 2;

QY 1 GCCTAGGAATGACAGGCTATCCACAGGCGAGCTGCATCCACCTT-GGCTGGGGGTGTCG 59
Db 84 GCCTAGGAATGACAGGCTATCCACAGGCGAGCTGCATCCACCTTGGGCTGGGGGTGTCG 143
QY 60 TCATTTGCTGCTATTAGAAAACGACAGGACCAATGCAATACCAACCGCTCCCGACTGTAA 119
Db 144 TCATTTGCTGCTATTAGAAAACGACAGGACCAATGCAATACCAACCGCTCCCGACTGTAA 203

```

```

QY 120 ACATAGGGATATGTTTCACTTAGCATGGACTTCTGGAGGGGCCAAGGAAGGCGGTC 179
Db 204 ACATAGGGATATGTTTCACTTAGCATGGACTTCTGGAGGGGCCAAGGAAGGCGGTC 263
QY 180 TGGAGTTTATTTAAATAGAGCAGTGTATTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 239
Db 264 TGGAGTTTATTTAAATAGAGCAGTGTATTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 323
QY 240 TCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 299
Db 324 TCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 383
QY 300 AAGTTTTCAGACTCAAGGCTATCAATGGAAATTTCAAGATCATCAAAATCAAAATCCACAG 359
Db 384 AAGTTTTCAGACTCAAGGCTATCAATGGAAATTTCAAGATCATCAAAATCAAAATCCACAG 443
QY 360 GGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
Db 444 GGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 503
QY 420 GGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
Db 504 GGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 563
QY 480 GGCATAAGAGGAGGAGATGCGACTCGAAGCAGAGAAATCGCTTACCGAGGAAATGAGC 539
Db 564 GGCATAAGAGGAGGAGATGCGACTCGAAGCAGAGAAATCGCTTACCGAGGAAATGAGC 623
QY 540 ATGATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 599
Db 624 ATGATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 683
QY 600 AGAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659
Db 684 AGAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 743
QY 660 ACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 719
Db 744 ACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 803
QY 720 ATGAGCGCGCTTCTGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
Db 804 ATGAGCGCGCTTCTGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 863
QY 780 AGGAGGCTCTGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 839
Db 864 AGGAGGCTCTGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 923
QY 840 TCCCAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 899
Db 924 TCCCAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 983
QY 900 AAGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959
Db 984 AAGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043
QY 960 ACCAGAGAGATGATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1019
Db 1044 ACCAGAGAGATGATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1103
QY 1020 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1079
Db 1104 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1148
QY 1080 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1139
Db 1149 ----- 1148
QY 1140 AGATCAGTTTCAAGAGAGGCTTAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1199
Db 1149 ----- 1148
QY 1200 CCCATCATGAAAGATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1259

```


Db 1149 ----- 1148
Qy 1260 TGAAGCAGAAAGAGAGAAAGAAATTAAGCCGAGCAGACAAAGAGATAGCAGATGAAC 1319
Db 1149 ----- 1148
Qy 1320 GAGCAGAATTGAAGCAGAAAGAAAGACAGCTGCCAAGAAAGAGAAAGGAGAGAGCAG 1379
Db 1149 ----- 1148
Qy 1380 AAGAGAGGAAAGGATGAGGAGGAAGAGAGAAAGGAGCAGAGGAGGAGAGGATGA 1439
Db 1149 ----- 1148
Qy 1440 AGGAGAGAGAAAGGAGCAGCAGAGAGGAGGAGAGGATTAAGGAGGAGAGAAAGGG 1499
Db 1149 ----- 1148
Qy 1500 CAGCAGAGGAGGAGCAGAGGATTAAGAGAGGAGAAAGGAGCAGAGAGGAGGCAAA 1559
Db 1149 ----- 1148
Qy 1550 GGGCCAGGCAGAGGAGGAGAGAGAGGCTTAAGTAGAGAGCAGAGACGTAACAGCAGC 1619
Db 1149 ----- 1148
Qy 1620 TAGAGAGAGAAACATGCCATGCAAGAGAGACAAAGATAAAAGGGGAAAGAGTAGACAGA 1679
Db 1149 ----- 1148
Qy 1680 AAATAGAGGGAATGGGTAAATGAAAGAGAAAGACAAAGAGATAAACTTCAGACAGCTG 1739
Db 1149 ----- 1148
Qy 1740 TCCTAAGAAACAGGAGAGAGAGAGGAGAACTTAAGTGCAAGCTAAAGAGAAAGGCTCC 1799
Db 1149 ----- 1148
Qy 1800 AAGAGACAAAGCCTACCTTCAAAAGAGAGAGAGATCAAGAGTAAAGATTAAAGAGACA 1859
Db 1149 ----- 1148
Qy 1860 AAGAACCCAAAGNAGAGATTAGAGCTTCATGATCGAAGAGAGGATTTACAGAGTTA 1919
Db 1179 AAGAACCCAAAGAGAAAGATTAGAGCTTCATGATCGAAGAGAGGATTTACAGAGTTA 1238
Qy 1920 AGTCGAGAGATGAGAGATTTCATGACCCACAAACTTAAACATACCTGAGAGATCTTCAGCC 1979
Db 1239 AGTCGAGAGATGAGAGATTTCATGACCCACAACTTAAACATACCTGAGAGATCTTCAGCC 1298
Qy 1980 GGCCTGAGAGGAGGCGCAGCTGGAACCAAGAGGCTGAGGCGGCCCCCGAGGTGGAAG 2039
Db 1299 GGCCTGAGAGGAGGCGCAGCTGGAACCAAGAGGCTGAGGCGGCCCCCGAGATGGAAG 1358
Qy 2040 CCGGCAAAAGGCTGAGAGGCTTCGTCGTCGCGGGAGACGAGAGCAAGATTGCG 2099
Db 1359 CCGGCAAAAGGCTGAGAGGCTTCGTCGTCGCGGGAGACGAGAGCCGAGAGCTGCG 1418
Qy 2100 AGAAGCTCAAAACAGAGCAGCAGAGGCGGCTTTGGAGCTGAGGAACTCAAGAAAAAGA 2159
Db 1419 AGAAGCTCAAAACAGAGCAGCAGAGGCGGCTTTGGAGCTGAGGAACTCAAGAAAAAGA 1478
Qy 2160 GGGAGAGAGAGAGGCTTCGAGGAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGCGG 2219
Db 1479 GGGAGAGAGAGAGGAGGCTTCGAGGAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGCGG 1538
Qy 2220 ATCGAAAACCTCAGAGAGGAGAGAGAGAGGAGGCTTAAAGAGAGAGATTGAAAGGCGAA 2279
Db 1539 ATCGAAAACCTCAGAGAGGAGAGAGAGAGAGGAGGCTTAAAGAGAGAGATTGAAAGGCGAA 1598
Qy 2280 GAGCAGAAGCTGCTGAGAAAGCCAGAGAGATGCCAGAGAGTGGCTTGTACATGACAGA 2339

Db 1599 GAGCAGAAGCTGCTGAGAAACGCCAGAGATGCCAGAGATGGCTTGTAGATGACAGA 1658
Qy 2340 AACCATTCAGTGTCTTCACTCTTAAGGTTTCATCTCTCAAGATAGAGAGGAGCAGAGAT 2399
Db 1659 AACCATTCAGTGTCTTCACTCTTAAGGTTTCATCTCTCAAGATAGAGAGGAGCAGAGAT 1718
Qy 2400 TTTTCAATTAAGTCTGTGAGAAAGCAGTGGTGTCAATCGACCATCAAGCAGCAATAG 2459
Db 1719 TTTTCAATTAAGTCTGTGAGAAAGCAGTGGTGTCAATCGACCATCAAGCAGCAATAG 1778
Qy 2460 TCTCAAGATTGACAGCAGACTGGAGCAGTATACCAAGTGAATGAGGGGAACAAAGAGCG 2519
Db 1779 TCTCAAGATTGACAGCAGACTGGAGCAGTATACCAAGTGAATGAGGGGAACAAAGAGCG 1838
Qy 2520 CAAAACCTTACAAAGCGGAGCCTCGGATCTTCTCTGTCTGTGAAAGTGTACGCAACA 2579
Db 1839 CAAAACCTTACAAAGCGGAGCCTCGGATCTTCTCTGTCTGTGAAAGTGTACGCAACA 1898
Qy 2580 TCAGAGTATGTTGGAGAAAGGAATGTGTTTATCTCCCACTGACAGCAGGACACCAA 2639
Db 1899 TCAGAGTATGTTGGAGAAAGGAATGTGTTTATCTCCCACTGACAGCAGGACACCAA 1958
Qy 2640 ATAAGGAACTGCTGGCTTGAAGTAGGGGTTTCTAGCCCATCAATGAATGGCTTACTA 2699
Db 1959 ATAAGGAACTGCTGGCTTGAAGTAGGGGTTTCTAGCCCATCAATGAATGGCTTACTA 2018
Qy 2700 AAAACCCAGATGGAACCAAGTCACTGCTGCCCAACCTTCTGACTTGAGACAGGAGAGC 2759
Db 2019 AAAACCCAGATGGAACCAAGTCACTGCTGCCCAACCTTCTGACTTGAGACAGGAGAGC 2078
Qy 2760 TATCAGCAGAGGAAACCTCTGGGAAAGCAATCTGTGGATAGGTCATCTCCCACTA 2819
Db 2079 TATCAGCAGAGGAAACCTCTGGGAAAGCAATCTGTGAAAGCAATCTGTGAAAGCAAT 2138
Qy 2820 AGGTTTGAGACAGTTCAGAAAGAACCCCAAGCTCAAGACGAGCAGCAGCTCAGTTGAG 2879
Db 2139 AGGTTTGAGACAGTTCAGAAAGAACCCCAAGCTCAAGACGAGCAGCAGCTCAGTTGAG 2198
Qy 2880 AGGCTAATTCGCTCGTGTGTTTGTATTTATGTTGATTTACTAAATGGTTCATTTCTTT 2939
Db 2199 AGGCTAATTCGCTCGTGTGTTTGTATTTATGTTGATTTACTAAATGGTTCATTTCTTT 2258
Qy 2940 TATTTTCAATATCCAGTAAACCCATGATATTTATCACTATATTTAATAATCACAGTCT 2999
Db 2259 TATTTTCAATATCCAGTAAACCCATGATATTTATCACTATATTTAATAATCACAGTCT 2318
Qy 3000 AGAGATGTTTCATGTTAAAGTACTGCTTTGCAAGAGGCTGTTCTTAAAGAAACCCAT 3059
Db 2319 AGAGATGTTTCATGTTAAAGTACTGCTTTGCAAGAGGCTGTTCTTAAAGAAACCCAT 2378
Qy 3060 GCTGTGAATAGAGACTTTTCTACTGATCATCACTCTGTATCTGAGCAGTGATACCA 3119
Db 2379 GCTGTGAATAGAGACTTTTCTACTGATCATCACTCTGTATCTGAGCAGTGATACCA 2438
Qy 3120 ACCACATCTGAAGTCAACAGAGATCCAAGTTTAAATTTGCTCGGAATGTGTCAGTA 3179
Db 2439 ACCACATCTGAAGTCAACAGAGATCCAAGTTTAAATTTGCTCGGAATGTGTCAGTA 2498
Qy 3180 TCTAGAAAAATGAACCGTAGTTTTGTTTTTTTAAATACAGAGTCACTGTTGTTTCTGCA 3239
Db 2499 TCTAGAAAAATGAACCGTAGTTTTGTTTTTTTAAATACAGAGTCACTGTTGTTTCTGCA 2558
Qy 3240 CTTTATAATAAGCATGGAAGAAATTTATCTTAGTAGGCAATTTGTAACACTTTTGTAAAGT 3299
Db 2559 CTTTATAATAAGCATGGAAGAAATTTATCTTAGTAGGCAATTTGTAACACTTTTGTAAAGT 2618
Qy 3300 AACCCATTTTCAGATTTGAAATACCTGCGATTAATGGTTCCTTTAAAGAAAAAGAAAGATG 3359
Db 2619 AACCCATTTTCAGATTTGAAATACCTGCAATTAATGGTTCCTTTAAAGAAAAAGAAAGATG 2678

RESULT 7

US-10-084-817-31

```

/ Sequence 31, Application US/10084817
/ Publication No. US20030119009A1
/ GENERAL INFORMATION:
/ APPLICANT: Susan Stuart
/ APPLICANT: Jed G. Nuchtern
/ APPLICANT: Sharon E. Plon
/ APPLICANT: Jason M. Shchet
/ TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
/ FILE REFERENCE: PA-0046 US
/ CURRENT APPLICATION NUMBER: US/10/084,817
/ CURRENT FILING DATE: 2002-02-25
/ PRIOR APPLICATION NUMBER: 60/270,784
/ PRIOR FILING DATE: 2001-02-23
/ NUMBER OF SEQ ID NOS: 365
/ SOFTWARE: PERL Program
/ SEQ ID NO 31
/ LENGTH: 4547
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. US20030119009A1 0224
/ US-10-084-817-31

```

Query Match.	48.4%;	Score 1793.6;	DB 14;	Length 4547;
Best Local Similarity	76.9%;	Prod. No. 0;		
Matches 2585;	Conservative 0;	Mismatches 9;	Indels 766;	Gaps 2;
Qy	1	GCCTAGGGAATGACAGGCATCTCCACAGCAGGTGCATCCACCTT-GCCTGGGGTGTGCG	59	
Db	84	GCTTAGGGAATGACAGGCATCTCCACAGCAGGTGCATCCACCTTGGCGTGGGGTGTGCG	143	
Qy	60	TCATTGGCTGCTATTGAAAAACGACAGGCAATGCAATCCACCGCCTCCCGACTGTAA	119	
Db	144	TCATTGGCTGCTATTGAAAAACGACAGGCAATGCAATCCACCGCCTCCCGACTGTGTAA	203	
Qy	120	ACATAGGGGATATGTTTCACTTAGCATGGACTTCTGGGAGGGGCAAGGAGGCGGCTC	179	
Db	204	ACATAGGGGATATGTTTCACTTAGCATGGACTTCTGGGAGGGGCAAGGAGGCGGCTC	263	
Qy	180	TGGAGTTTATTGAAATAGAGCAGTGTGTTATTCGGCTGCTGCTGCTGCCCGCTGCTTGCTC	239	
Db	264	TGGAGTTTATTGAAATAGAGCAGTGTGTTATTCGGCTGCTGCTGCTGCCCGCTGCTTGCTC	323	
Qy	240	TCTGCTGTGCTCCTGCTTTAAGAAATCAGTCCCTTCCCTTCCGACTTGGTCTCCGGGAAG	299	
Db	324	TCTGCTGTGCTCCTGCTTTAAGAAATCAGTCCCTTCCCTTCCGACTTGGTCTCCGGGAAG	383	
Qy	300	AAGTTTTCAGCTACAAGGTATCATTTGGAAACATTTCAAGATCATCAAAATTCAAATTCACAG	359	
Db	384	AAGTTTTCAGCTACAAGGTATCATTTGGAAACATTTCAAGATCATCAAAATTCAAATTCACAG	443	
Qy	360	GGATTGGTGACCAACGACAGAGGCTCAGACATCTGATTGCTGACCTGTCCAGACATCATCT	419	
Db	444	GGATTGGTGACCAACGACAGAGGCTCAGACATCTGATTGCTGACCTGTCCAGACATCATCT	503	
Qy	420	GGTCTCCCTGAAACCTGAAATCAACCATCGATGATTTTGAGCGTTCGACAGAACTTAGAA	479	
Db	504	GGTCTCCCTGAAACCTGAAATCAACCATCGATGATTTTGAGCGTTCGACAGAACTTAGAA	563	
Qy	480	GGCAAGAGGGAGGAGATGCGACTTCGACGAGAAAGAAATCGCTTACCAGAGGAATGACG	539	
Db	564	GGCAAGAGGGAGGAGATGCGACTTCGACGAGAAAGAAATTCGCTTACCAGAGGAATGACG	623	
Qy	540	ATGATGAAGAGAGGCAGCCCGGAAACGGCGCGCGAGCCCGACAGGAACCGGCTCGCGC	599	
Db	624	ATGATGAAGAGAGGCAGCCCGGAAACGGCGCGCGAGCCCGACAGGAACCGGCTCGCGC	683	
Qy	600	AGACGAGGAGGAAGAAATCTTTGGGACAGGTGACCGACACGAGTGGAGGTGAATGCCACA	659	
Db	684	AGAAAGCAGGAGGAAGAAATCTTTGGGACAGGTGACCGACACGAGTGGAGGTGAATGCCACA	743	
Qy	660	ACAGTGTGCTGACGAGGAGGCGCAAGACAAACCAACCAACACTCAAGTGGAGAGGGATG	719	

Db	1149	-----	1149	
Qy	1900	AAGAAGACAAGCCTACCTTCAAAAAAAGAGAGATCAAAAGATGAAAAAGATTAATAAAGAGACA	1859	
Db	1149	-----AGATCAAAAGATGAAAAGATTAATAAAGAGACA	1178	
Qy	1960	AAGAACCCAAAGAGAGTTAGAGCTTTCATGGATCGAAAGAGGGATTTACAGAGTTA	1919	
Db	1179	AAGAACCCAAAGAGAGTTAAGAGCTTTCATGGATCGAAAGAGGGATTTACAGAGTTA	1238	
Qy	1920	AGTCGAGAATGGAGAAATTCATGACCCACAAAATTAAACATACCTGAGAAATCTTTGAGCC	1979	
Db	1239	AGTCGAGAATGGAGAAATTCATGACCCACAAAATTAAACATACCTGAGAAATCTTTGAGCC	1298	
Qy	1980	GCCTGGAGGGAGGCCAGGTCGACACCAAGAGGCTGAGGGCCGCCCCAGGTTGGAG	2039	
Db	1299	GCCTGGAGGGAGGCCAGGTCGACACCAAGAGGCTGAGGGCCGCCCCAGATGGAG	1358	
Qy	2040	CCGGCAAAAGGCTGGAGGAGCTTCGTGCTCGCGGGGAGACCGAGAGCGAAGAGTTCG	2099	
Db	1359	CCGGCAAAAGGCTGGAGGAGCTTCGTGCTCGCGGGGAGACCGAGAGCGAAGAGTTCG	1418	
Qy	2100	AGAAGCTCAAAAGAGACGAGAGCGCGCTTTGGAGCTGGAGAACTCAAGAAAAAGA	2159	
Db	1419	AGAAGCTCAAAAGAGACGAGAGCGCGCTTTGGAGCTGGAGAACTCAAGAAAAAGA	1478	
Qy	2160	GGGAGGAGAGAGGAGTTCCTCGAGGAGAGAGCAGAGGAGGAGCAGGAGGAGCGC	2219	
Db	1479	GGGAGGAGAGAGGAGTTCCTCGAGGAGAGAGCAGAGGAGGAGCAGGAGGAGCGC	1538	
Qy	2220	ATCGAAAACTCAGAGAGGAGAAAGAGAGGAGGCTTAAAGAAAGAGATTTGAAAGCGGAA	2279	
Db	1539	ATCGAAAACTCAGAGAGGAGAAAGAGAGGAGGCTTAAAGAAAGAGATTTGAAAGCGGAA	1598	
Qy	2280	GAGCAGAGCTGCTGAGAAACGCCAGAGATGCCAGAGAGTGGCTTGTTCAGATGACAGA	2339	
Db	1599	GAGCAGAGCTGCTGAGAAACGCCAGAGATGCCAGAGAGTGGCTTGTTCAGATGACAGA	1658	
Qy	2340	AACCATTTCAAGTGTTTTCACTCTCTAAAGGTTTCATCTCTCAAGATAGAAGCGCAGAGAT	2399	
Db	1659	AACCATTTCAAGTGTTTTCACTCTCTAAAGGTTTCATCTCTCAAGATAGAAGCGCAGAGAT	1718	
Qy	2400	TTTTGAATAGTCTGTGAGAAAGAGTGTTGTCMAATCGACCCATCAAGCAGCAGATAG	2459	
Db	1719	TTTTGAATAGTCTGTGAGAAAGAGTGTTGTCMAATCGACCCATCAAGCAGCAGATAG	1778	
Qy	2460	TCTCAAGATTGACAGCAGACTGGAGCAGATATACGAGTCAATTCAGGGAAACAAAAGCG	2519	
Db	1779	TCTCAAGATTGACAGCAGACTGGAGCAGATATACGAGTCAATTCAGGGAAACAAAAGCG	1838	
Qy	2520	CAAAACCTTCAAAAGCCGGCAGCCTCGATCTTCTGTTCTCTGCTGAGAGTGTCGCAACA	2579	
Db	1839	CAAAACCTTCAAAAGCCGGCAGCCTCGATCTTCTGTTCTCTGCTGAGAGTGTCGCAACA	1898	
Qy	2580	TCAAGAGTATGGGAGAAAGGGAATGTGTTTTTCATCCCCCATCTCAGCAGGCACACCAA	2639	
Db	1899	TCAAGAGTATGGGAGAAAGGGAATGTGTTTTTCATCCCCCATCTCAGCAGGCACACCAA	1958	
Qy	2640	ATAAGAAACTGCTGGCTTGAAGGTAGGGTTTTCTAGCCGATCAATGAATGGCTAACTA	2699	
Db	1959	ATAAGAAACTGCTGGCTTGAAGGTAGGGTTTTCTAGCCGATCAATGAATGGCTAACTA	2018	
Qy	2700	AAACCCAGATGGAACAAGTCACTGCTCCCAACCTTCTGACTTTGAGACCAGAGAGCG	2759	
Db	2019	AAACCCAGATGGAACAAGTCACTGCTCCCAACCTTCTGACTTTGAGACCAGAGAGCG	2078	
Qy	2760	TATTCAGCAAGCGGAACCTCTGGGAAAAGCAATCTGTGGATAAGGTCACTTCCCCCACTA	2819	
Db	2079	TATTCAGCAAGCGGAACCTCTGGGAAAAGCAATCTGTGGATAAGGTCACTTCCCCCACTA	2138	
Qy	2820	AGGTTTGACAGTTTCAGAAAAGAACCAAGCTCAAGACGACAGACGAGCTCAGTTGTAG	2879	
Db	2139	AGGTTTGACAGTTTCAGAAAAGAACCAAGCTCAAGACGACAGACGAGCTCAGTTGTAG	2198	

2880	Qy	AGGCGTAATTCGCTCGTGTGTTTGATTTATGTTGATTTACTAAATTCGGTTTCATATCTTT	2939
2199	Db	AGGCGTAATTCGCTCGTGTGTTTGATTTATGTTGATTTACTAAATTCGGTTTCATATCTTT	2558
2940	Qy	TATTTTTCATATCCCAAGTAAACCCCATGTATATATACACTATATATTTAAATAATCACAGTCT	2999
2259	Db	TATTTTTCATATCCCAAGTAAACCCCATGTATATATACACTATATTTAAATAATCACAGTCT	2318
3000	Qy	AGAGATGTTTCATCGTAAAAAGTACTGCGTTTGCACAGAGCGCTGTTTCTAAAGAAACCCAT	3059
2319	Db	AGAGATGTTTCATCGTAAAAAGTACTGCGTTTGCACAGAGCGCTGTTTCTAAAGAAACCCAT	2378
3060	Qy	GCTGTGAAATAGAGACTTTTCTACTCATCATATAACTCTGTATCTCGACAGATGATACCA	3119
2379	Db	GCTGTGAAATAGAGACTTTTCTACTCATCATATAACTCTGTATCTCGACAGATGATACCA	2438
3120	Qy	ACCACATCTGAAGTCAACAGAAGATCCAAGTTTAAATTCGCTCGCGAATGTGTGCAGTA	3179
2439	Db	ACCACATCTGAAGTCAACAGAAGATCCAAGTTTAAATTCGCTCGCGAATGTGTGCAGTA	2498
3180	Qy	TCTAGAAAAATGAACCGTAGTTTTGTGTTTTTTTAAATACAGAAGTCAATGTGTTTCTGCA	3239
2499	Db	TCTAGAAAAATGAACCGTAGTTTTGTGTTTTTTTAAATACAGAAGTCAATGTGTTTCTGCA	2558
3240	Qy	CTTTTATAATAAAGCATGGAAGAAATATCTTAGTAGGCAATTTGTAAACACTTTTTTGAAAGT	3299
2559	Db	CTTTTATAATAAAGCATGGAAGAAATATCTTAGTAGGCAATTTGTAAACACTTTTTTGAAAGT	2618
3300	Qy	AACCCATTTTCAGATTTGAAATACCTGGATTAATCGGTGCTTTTAAAAAAAAGAAAAAGATG	3359
2619	Db	AACCCATTTTCAGATTTGAAATACCTGGATTAATCGGTGCTTTTAAAAAAAAGAAAAAGATG	2678

RESIST. T 8

```

RES001 68
US-10-062-674-1816
; Sequence 1816, Application US/10062674
; Publication No. US20040005559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 09/625,102
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 2217
; SOFTWARE: PERL Program
; SEQ ID NO 1816
; LENGTH: 4261
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: incyte ID No. US20040005559A1 256147.57
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (4261)
; OTHER INFORMATION: a, t, c, g, or other
US-10-062-674-1816

```

Query Match	47.8%	Score 1771;	DB 15;	Length 4263;
Best Local Similarity	76.7%;	Pred. No. 0;		
Matches 2578:	Conservative	0:	Mismatches	16:
				Indels 768:
				Gaps 3:

1 GCCTAGGGAAATGACAGGCATCTCCACAGGACAGCTGCATCCACCTT--GGCTGGGGTCTCG 59
84 GCCTAGGGAAATGACAGGCATCTCCACAGGACAGCTGCATCCACCTTGGGCTGGGGTCTCG 143
60 TCATTGGCTGCCATTAGAAAAACGACAGACAAATGCTATCCACCGCCTCCCGACTGAA 119
144 TCATTGGCTGCCATTAGAAAAACGACAGACAAATGCTATCCACCGCCTCCCGACTGAA 203

Db 849 ----- 848
 QY 1597 AGAGCAGAAAGCTAACAAGCAGTAGAAGAGAGAAAAAATCATGCCATGCAAGACACAAAGAT 1656
 Db 849 ----- 848
 QY 1657 AAAAGGGGAAAGGTAGAACACAAATAATAGAGGGAAATGGGTAAATGAAAGAAAGCACA 1716
 Db 849 ----- 848
 QY 1717 AGAAGATAAACTTCAGACAGCTGTCTCTAAAGAAACAGGGAGAGAGAGGAACCTAAAGT 1776
 Db 849 ----- 848
 QY 1777 GCAAGCTAAAGAGAAAGCTCCAAAGAGACAAAGCTACCTTCAAAAAAGAGAGATCAA 1836
 Db 849 ----- 855
 QY 1837 AGATGAAAGATTAAAAAGGACAAAGAACCCCAAGAAAGTTAAGAGCTTCATGATCG 1896
 Db 856 AGATGAAAGATTAAAAAGGACAAAGAACCCCAAGAGAGTTAAGAGCTTCATGATCG 915
 QY 1897 AAAGAGGGATTTACAGAAATTAAAGTCGAGAAATGAGAGAAATTCATGACCCACAACTTAA 1956
 Db 916 AAAGAGGGATTTACAGAAATTAAAGTCGAGAAATGAGAGAAATTCATGACCCACAACTTAA 975
 QY 1957 ACATACTGAGAAATCTTTTCAGCGCGCTGAGGAGGGCCAGCGTGGACACCAAGGAGGC 2016
 Db 976 ACATACTGAGAAATCTTTTCAGCGCGCTGAGGAGGGCCAGCGTGGACACCAAGGAGGC 1035
 QY 2017 TGAGGGGCCCCCAGGTGGAAGCGGCAAAAGGCTGGAAGAGCTTCGTCGTCGCGG 2076
 Db 1036 TGAGGGGCCCCCAGGTGGAAGCGGCAAAAGGCTGGAAGAGCTTCGTCGTCGCGG 1095
 QY 2077 GGAGACGAGAGCGAAGAGTTTCGAGAGCTCAAAAGAGAGCAGAGGAGGCGCTTTGGA 2136
 Db 1096 GGAGACGAGAGCGAAGAGTTTCGAGAGCTCAAAAGAGAGCAGAGGAGGCGCTTTGGA 1155
 QY 2137 GCTGGAGAACTCAAGAAAGAGGAGGAGAGAGAGAGAGAGTCTCTGGAGAGAGAGCA 2196
 Db 1156 GCTGGAGAACTCAAGAAAGAGGAGGAGAGAGAGAGAGTCTCTGGAGAGAGAGCA 1215
 QY 2197 GAGGAGGAGCAGAGGAGGAGCGATCGAAAGCTCAAGAGAGGAGAGAGAGAGAGGAGGCT 2256
 Db 1216 GAGGAGGAGCAGAGGAGGAGCGATCGAAAGCTCAAGAGAGGAGAGAGAGAGGAGGCT 1275
 QY 2257 AAAAGAGAGATTGAAAGCGGAGAGCAGAGCTGCTGAGAAAGCGCAGAGATGCCAGA 2316
 Db 1276 AAAAGAGAGATTGAAAGCGGAGAGCAGAGCTGCTGAGAAAGCGCAGAGATGCCAGA 1335
 QY 2317 AGATGGCTTGCAGATGACAAAGAAACCATTCAGTGTCTTCACTCCTTAAAGGTTTCATCTCT 2376
 Db 1336 AGATGGCTTGCAGATGACAAAGAAACCATTCAGTGTCTTCACTCCTTAAAGGTTTCATCTCT 1395
 QY 2377 CAAGATAGAGAGCGAGCAGAAATTTTGAATAAGTCTGTGCAGAAAGCAGTGGTGTCAA 2436
 Db 1396 CAAGATAGAGAGCGAGCAGAAATTTTGAATAAGTCTGTGCAGAAAGCAGTGGTGTCAA 1455
 QY 2437 ATCGACCCATCAAGCAGCAGATAGTCTCCAGATTGACAGCAGAGCTGAGCAGTATACCAG 2496
 Db 1456 ATCGACCCATCAAGCAGCAGATAGTCTCCAGATTGACAGCAGAGCTGAGCAGTATACCAG 1515
 QY 2497 TGCAATTGAGGGAACAAAAGCGCAAAACCTACAAAGCGGCGCTCGGATCTTCTGT 2556
 Db 1516 TGCAATTGAGGGAACAAAAGCGCAAAACCTACAAAGCGGCGCTCGGATCTTCTGT 1575
 QY 2557 TCCTGCTGAAGTGTACCAACATCAGAGTATGTGGGAGAAAGGAGATGTGTTTCATC 2616
 Db 1576 TCCTGCTGAAGTGTACCAACATCAGAGTATGTGGGAGAAAGGAGATGTGTTTCATC 1635
 QY 2617 CCCCCTGAGGAGGACACCAACCAATAGGAACTGCTGGCTTGAAGTACGGGTTCTAG 2676
 Db 1636 CCCCCTGAGGAGGACACCAACCAATAGGAACTGCTGGCTTGAAGTACGGGTTCTAG 1695

QY 2677 CCGCATCAATGAATGGCTAATAAAACCCAGATGGAACAAAGTCACTGCTCCCAACC 2736
 Db 1696 CCGCATCAATGAATGGCTAATAAAACCCAGATGGAACAAAGTCACTGCTCCCAACC 1755
 QY 2737 TTCTGACTTCAGACACGAGAGCTATCCAGCAAGCGGAACCTCTGGGAAAAGCAATCTGT 2796
 Db 1756 TTCTGACTTCAGACACGAGAGCTATCCAGCAAGCGGAACCTCTGGGAAAAGCAATCTGT 1815
 QY 2797 GGATAAGGTCACTTCCCCCACTAAGGTTTGACAGAGTTCCAGAAAGAACCCAGCTCAAG 2856
 Db 1816 GGATAAGGTCACTTCCCCCACTAAGGTTTGACAGAGTTCCAGAAAGAACCCAGCTCAAG 1875
 QY 2857 ACGCAGACGAGCTCAGTTGTAGAGGCTAATTCGCTCTGTTGTTATTTATGTTGATTT 2916
 Db 1876 ACGCAGACGAGCTCAGTTGTAGAGGCTAATTCGCTCTGTTGTTATTTATGTTGATTT 1935
 QY 2917 ACTAAATGGGTTCAATATCTTTTATTTTCAATATCCAGTAACCCATGTTATATC 2976
 Db 1936 ACTAAATGGGTTCAATATCTTTTATTTTCAATATCCAGTAACCCATGTTATATC 1995
 QY 2977 ACTATATTTAATAATCACAGTCTAGAGATGTTTCATGGTAAAGTACTGCTTTGCACAGG 3036
 Db 1996 ACTATATTTAATAATCACAGTCTAGAGATGTTTCATGGTAAAGTACTGCTTTGCACAGG 2055
 QY 3037 AGCTGTTTCTAAAGAAACCCATGCTGAAATAGAGACTTTTCTACTGATCATTAAC 3096
 Db 2056 ATCTGTTTCTAAAGAAACCCATGCTGAAATAGAGACTTTTCTACTGATCATTAAC 2115
 QY 3097 TCTGATCTGAGCAGTATACCAACACATCTGAAAGTCAACAGAGATCCAAGTTTAAA 3156
 Db 2116 TCTGATCTGAGCAGTATACCAACACATCTGAAAGTCAACAGAGATCCAAGTTTAAA 2175
 QY 3157 TTGCTCGGAAATGTGTCAGATCTAGAAATAGAACCGTAGTTTGTGTTTAAAT 3216
 Db 2176 TTG-CTGCGAATGTGTCAGATCTAGAAATAGAACCGTAGTTTGTGTTTAAAT 2234
 QY 3217 ACAGAGTCATGTTGTTTCTGCACCTTATAAAGCATGGAAGAAATATCTTAGTAGG 3276
 Db 2235 ACAGAGTCATGTTGTTTCTGCACCTTATAAAGCATGGAAGAAATATCTTAGTAGG 2294
 QY 3277 CAATTGTAACCTTTTGAAGTAAACCCATTTGAAATAGTGAATATCTGCAATATGTTG 3336
 Db 2295 CAATTGTAACCTTTTGAAGTAAACCCATTTGAAATAGTGAATATCTGCAATATGTTG 2354
 QY 3337 CTTTAAAAAAGAAAGAGATGCGCGACAGTATTCTTT 3378
 Db 2355 TCTTTAAAAAAGAAAGATGTTTAAAGTATTACTTTT 2396

RESULT 11

US-09-765-231A-32
 ; Sequence 32, Application US/09765231A
 ; Patent No. US20020119452A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Searle/Monsanto
 ; APPLICANT: Phippard, Deborah
 ; APPLICANT: Vasanthakumar, Geetha
 ; APPLICANT: Dotson, Stanton
 ; APPLICANT: Ma, Xiao-Jun
 ; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,
 ; FILE OF INVENTION: vectors, and cells
 ; FILE REFERENCE: SO-3221 PR
 ; CURRENT APPLICATION NUMBER: US/09/765,231A
 ; CURRENT FILING DATE: 2001-01-18
 ; NUMBER OF SEQ ID NOS: 82
 ; SEQ ID NO 32
 ; LENGTH: 3342
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-765-231A-32

Query Match 39.3%; Score 1455.8; DB 9; Length 3342;

Best Local Similarity 99.4%; Pred. No. 0; Matches 1482; Conservative 0; Mismatches 7; Indels 2; Gaps 2;									
QY	1871	GAAGAAGTTAAGAGCTTCTGAGATCGAAGAGGATTTACAGAGTTAAAGTCGAGAAAT	1930						
DB	1	GAAGAAGTTAAGAGCTTCTGAGATCGAAGAGGATTTACAGAGTTAAAGTCGAGAAAT	60						
QY	1931	GGAGAATTCATGACCCACAACTTAACATACACTGAGAAATCTTTTCAGCGCCCTGGAGG	1990						
DB	61	GGAGAATTCATGACCCACAACTTAACATACACTGAGAAATCTTTTCAGCGCCCTGGAGG	120						
QY	1991	AGGGCAGCGTGCACACCAAGGAGCTGAGGGCCGCCAGGTGGAGCGGCAAAAGG	2050						
DB	121	AGGGCAGCGTGCACACCAAGGAGCTGAGGGCCGCCAGGTGGAGCGGCAAAAGG	180						
QY	2051	CTGGAGAGCTTCGTCTGCTCGCGGGAGACCGAGAGCGAAGAGTTTCGAGAACTCAA	2110						
DB	181	CTGGAGAGCTTCGTCTGCTCGCGGGAGACCGAGAGCGAAGAGTTTCGAGAACTCAA	240						
QY	2111	CAGAACGACGAGGCGGCTTTGGAGCTGGAGAACTCAAGAAACAGAGGAGGAGAGA	2170						
DB	241	CAGAACGACGAGGCGGCTTTGGAGCTGGAGAACTCAAGAAACAGAGGAGGAGAGA	300						
QY	2171	AGGAAGTCTCTGAGGAGGAGAGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2230						
DB	301	AGGAAGTCTCTGAGGAGGAGAGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG	360						
QY	2231	AGAGAGGAGAGAGAGAGAGGAGGCTTAAGAGAGAGATTGAAGCGAGAGCAGAGCT	2290						
DB	361	AGAGAGGAGAGAGAGAGAGGAGGCTTAAGAGAGAGATTGAAGCGAGAGCAGAGCT	420						
QY	2291	GCTGAGAAACGCCAGAGAGATGCCAGAGATGGCTTGTGAGATGACAGAAACCAATTC	2350						
DB	421	GCTGAGAAACGCCAGAGATGCCAGAGATGGCTTGTGAGATGACAGAAACCAATTC	480						
QY	2351	TGTTTCATCTCTAAGAGTCTCATCTCTCAAGATGAGAGAGGAGAGAGATTTTGAAT	2410						
DB	481	TGTTTCATCTCTAAGAGTCTCATCTCTCAAGATGAGAGAGGAGAGATTTTGAAT	540						
QY	2411	TCGTGCAGAAAGACGCTGTGTCAAACTCGACCATCAAGCAGCAATAGTCTCAAGAT	2470						
DB	541	TCGTGCAGAAAGACGCTGTGTCAAACTCGACCATCAAGCAGCAATAGTCTCAAGAT	600						
QY	2471	GACAGCAGCTGGAGCAGTATACCAAGTGCATTTGAGGGAACAAAGCGCAAAACCTCA	2530						
DB	601	GACAGCAGCTGGAGCAGTATACCAAGTGCATTTGAGGGAACAAAGCGCAAAACCTCA	660						
QY	2531	AGCGCGCAGCTCGGATCTCTCTGCTGAGGTGTACCAACATCAAGAGTATG	2590						
DB	661	AGCGCGCAGCTCGGATCTCTCTGCTGAGGTGTACCAACATCAAGAGTATG	720						
QY	2591	TGGGAGAAAGGGAATGTGTTTTCATCCCCCACTGCAGCAGGCACACCAAAATAGGAACT	2650						
DB	721	TGGGAGAAAGGGAATGTGTTTTCATCCCCCACTGCAGCAGGCACACCAAAATAGGAACT	780						
QY	2651	G-CTGGCTTGAAGTAGGGTTCTAGCCGATCAATGAATGGCTTAATAAACCACCA	2709						
DB	781	GCTGGCTTGAAGTAGGGTTCTAGCCGATCAATGAATGGCTTAATAAACCACCA	840						
QY	2710	TGGAAACAAGTCACTGCTCCCAACCTTCTGACTTGAGACGAGGAGAGTATCCAGCA	2769						
DB	841	TGGAAACAAGTCACTGCTCCCAACCTTCTGACTTGAGACGAGGAGAGTATCCAGCA	900						
QY	2770	GCAGAACCTTCGGAAAGCAATCTGGATGAAGGTCACTTCGCCCACTAAGTTTGA	2828						
DB	901	GCAGAACCTTCGGAAAGCAATCTGGATGAAGGTCACTTCGCCCACTAAGTTTGA	960						
QY	2829	ACAGTCCAGAAAGACCCAGAGCTCAAGCAGGAGGAGGAGGAGGAGGAGGAGG	2888						
DB	961	ACAGTCCAGAAAGACCCAGAGCTCAAGCAGGAGGAGGAGGAGGAGGAGGAGG	1020						
QY	2889	TCGCTCTGTTTGTGATTTATGTTTCTGATTTTAAATTTGGGTTTCATTTATTTTCA	2948						

RESULT 12
US-09-925-301-100
; Sequence 100, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 100
; LENGTH: 2511
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (12)-
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (28)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (44)-
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (2456)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (2488)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (2511)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-100
Query Match 15.8%; Score 594.2; DB 9; Length 2511;

NAME/KEY: misc_feature	15.8%;	Score 584.2;	DB 14;	Length 2511;
LOCATION: (44)..(44)	Best Local Similarity	98.7%;	Pred: No. 1.8e-138;	
OTHER INFORMATION: n equals a,t,g, or c	Matches 569;	Conservative 0;	Mismatches 8;	Indels 0;
NAME/KEY: misc_feature				Gaps 0;
LOCATION: (2456)..(2456)				
OTHER INFORMATION: n equals a,t,g, or c				
NAME/KEY: misc_feature				
LOCATION: (2488)..(2488)				
OTHER INFORMATION: n equals a,t,g, or c				
NAME/KEY: misc_feature				
LOCATION: (2511)..(2511)				
OTHER INFORMATION: n equals a,t,g, or c				
US-10-106-698-183				
Query Match				
Best Local Similarity				
Matches 569;				
Conservative 0;				
Mismatches 8;				
Indels 0;				
Gaps 0;				
2763	CCAGCAAGCGGAACCTCTCGGAAAAGCAATCTGTGGATAAGTCACTTCCCGCACTAAGG	2822		
73	CGAGCAAGCGGAACCTCTCGGAAAAGCAATCTGTGGATAAGTCACTTCCCGCACTAAGG	132		
2823	TTTGAGACAGTTCCAGAAAGACCCAGCTCAAGACGCGAGGACGAGTCACTAGTTGTAGAGG	2882		
133	TTTGAGACAGTTCCAGAAAGACCCAGCTCAAGACGCGAGGACGAGTCACTAGTTGTAGAGG	192		
2883	GCTAAATCCGCTCTGTTTCTATTATGTATTGTTTAAATCTAAATGGGTTTCATTCTTTTAT	2942		
193	GCTAAATCCGCTCTGTTTCTATTATGTATTGTTTAAATCTAAATGGGTTTCATTCTTTTAT	252		
2943	TTTTCAATATCCAGTAACCCATGTATATATCACTATATTTAATAACACAGTCTAGA	3002		
253	TTTTCAATATCCAGTAACCCATGTATATATCACTATATTTAATAACACAGTCTAGA	312		
3003	GATGTTTCATGTTAAAGTACTGCGTTTGCACAGAGGCGCTCTTCTAAAGAAACCCATGCT	3062		
313	GATGTTTCATGTTAAAGTACTGCGTTTGCACAGAGGCGCTCTTCTAAAGAAACCCATGCT	372		
3063	GTCAATATAGAGACTTTTCTACTGATCATCATAACTCTGTATCTGAGCAGTATACCAACC	3122		
373	GTCAATATAGAGACTTTTCTACTGATCATCATAACTCTGTATCTGAGCAGTATACCAACC	432		
3123	ACATCTGAAGTCAACAGAGATCCCAAGTTTAAATTCCTCGGAATGTGTGCAAGTATCT	3182		
433	ACATCTGAAGTCAACAGAGATCCCAAGTTTAAATTCCTCGGAATGTGTGCAAGTATCT	492		
3183	AGAAAAATGAACCGTAGTTTTTGTGTTTTTAAATACAGAAAGTCACTGTTGTCGCACTT	3242		
493	AGAAAAATGAACCGTAGTTTTTGTGTTTTTAAATACAGAAAGTCACTGTTGTCGCACTT	552		
3243	TATAATAAGCATGGAAGAAATTTATCTTAGTAGGCAATTCGTACACCTTTTGAAGTAAC	3302		
553	TATAATAAGCATGGAAGAAATTTATCTTAGTAGGCAATTCGTACACCTTTTGAAGTAAC	612		
3303	CCATTTTCAGATTTGAAATCTCGCGATAATCGTTGCTTTTAAAAAAGAAAAAGATG	3359		
613	CCATTTTCAGATTTGAAATCTCGCGATAATCGTTGCTTTTAAAAAAGAAAAAGATG	669		
RESULT 14				
US-09-974-298-35				
Sequence 35, Application US/09974298				
Patent No. US20020156263A1				
GENERAL INFORMATION:				
APPLICANT: Chen, Hwei-Mei				
TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER				
FILE REFERENCE: PA-0037 P				
CURRENT APPLICATION NUMBER: US/09/974,298				
CURRENT FILING DATE: 2001-10-04				
PRIOR APPLICATION NUMBER: 60/238,331				
PRIOR FILING DATE: 2000-05-10				
NUMBER OF SEQ ID NOS: 194				
SOFTWARE: PERL Program				

; SEQ ID NO 35
 ; LENGTH: 2209
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20020156263A1 022404.23
 US-09-974-298-35

Query Match 14.8%; Score 546.8; DB 9; Length 2209;
 Best Local Similarity 78.3%; Pred. No. 6.2e-129;
 Matches 808; Conservative 0; Mismatches 7; Indels 217; Gaps 4;

QY	2328	CAGATGCAAGAACCACTCAAGTGTTCCTCAAGAGGTTCACTCTCAAGATAGAAG	2387
DB	1	CAGATGCAAGAACCACTCAAGTGTTCCTCAAGAGGTTCACTCTCAAGATAGAAG	60
QY	2388	AGCGAGCAGAAATTTTGAATAGTCTGTGCAGAAAGCAGTGTGTCAAATCGACCCATC	2447
DB	61	AGCGAGCAGAAATTTTGAATAGTCTGTGCAGAAAGCAGTGTGTCAAATCGACCCATC	120
QY	2448	AAGCAGCAATAGTCTCAAGATGACAGCAGACTGAGCAGTATACAGTGCATTTGAGG	2507
DB	121	AAGCAGCAATAGTCTCAAGATGACAGCAGACTGAGCAGTATACAGTGCATTTGAGG	180
QY	2508	GACAAAAGCGCAAACTCAAGAGCGGCGCTCGGATCTTCTGTTCTGCTGAAG	2567
DB	181	GACAAAAGCGCAAACTCAAGAGCGGCGCTCGGATCTTCTGTTCTGCTGAAG	240
QY	2568	GTGTACGCAACATCAAGATGATGTGGGAGAAAGGGAATGTGTTTCAATCCGCTGCG	2627
DB	241	GTGTACGCAACATCAAGATGATGTGGGAGAAAGGGAATGTGTTTCAATCCGCTGCG	300
QY	2628	CAGCAGCAGCAATCAAGATGATGTGGGAGAAAGGGAATGTGTTTCAATCCGCTGCG	2687
DB	301	CAGCAGCAGCAATCAAGATGATGTGGGAGAAAGGGAATGTGTTTCAATCCGCTGCG	360
QY	2688	AATGGCTAACTAAACCCAGATGAAACCACTGCTCTCCCAACCTTCTGACTTGA	2747
DB	361	AATGGCTAACTAAACCCAGATGAAACCACTGCTCTCCCAACCTTCTGACTTGA	420
QY	2748	GACGAGGAGCTATCCAGCAGCGGACCTCTGGGAAAGCAATCTGTGGATAAGTCA	2807
DB	421	GACGAGGAGCTATCCAGCAGCGGACCTCTGGGAAAGCAATCTGTGGATAAGTCA	480
QY	2808	CTTCCGCTAAAGGTTTGAGACAGTCTCAGAAAGAACCCAGCTCAAGACGAGCAGCA	2867
DB	481	CTTCCGCTAAAGGTTTGAGACAGTCTCAGAAAGAACCCAGCTCAAGACGAGCAGCA	492
QY	2868	GCTCAGTTGTAGAGGGCTAAATCGCTCTGTTTGTATTTATGTTGATTTACTAAATGGG	2927
DB	493	GCTCAGTTGTAGAGGGCTAAATCGCTCTGTTTGTATTTATGTTGATTTACTAAATGGG	492
QY	2928	TTCAATATCTTTTATTTTCAATATCCAGTAAACCCATGTATATTACATATATTTAA	2987
DB	493	TTCAATATCTTTTATTTTCAATATCCAGTAAACCCATGTATATTACATATATTTAA	492
QY	2988	TAAATCAGTCTAGAGATGTTCAATGGTAAAGTACTGCTTTGCAAGGAGCTGTTTCT	3047
DB	493	TAAATCAGTCTAGAGATGTTCAATGGTAAAGTACTGCTTTGCAAGGAGCTGTTTCT	506
QY	3048	AAAGAAACCCATGCTGGAATAGAGACTTTTCTACTGATCATCATAACTCTGTATCTGA	3107
DB	507	AAAGAAACCCATGCTGGAATAGAGACTTTTCTACTGATCATCATAACTCTGTATCTGA	566
QY	3108	GCAGTGATACCAACCACTCTGAAGTCAACAGAGATCCAAAGTTTAAATTTGCTCGGA	3167
DB	567	GCAGTGATACCAACCACTCTGAAGTCAACAGAGATCCAAAGTTTAAATTTGCTCGGA	626
QY	3168	ATGTGTGAGTATCTAGAAATGACCGTAGTTTGTGTTTGTATTTTAAATACAGAGTCAT	3227
DB	627	ATGTGTGAGTATCTAGAAATGACCGTAGTTTGTGTTTGTATTTTAAATACAGAGTCAT	686

QY	3228	GTCTTTCTGCACCTTTATAATAAAGCATGGAAGAAATTTCTTTAGTAGGCAATTTGAACA	3287
DB	687	GTCTTTCTGCACCTTTATAATAAAGCATGGAAGAAATTTCTTTAGTAGGCAATTTGAACA	745
QY	3288	CTTTTGAAGTAACCCATTTTCAGATTTTGAATCTGCGATATGTTGCTTTAAAAAAA	3347
DB	746	CTTTTGAAGTAACCCATTTTCAGATTTTGAATCTGCGATATGTTGCTTTAAAAAAA	803
QY	3348	AAGAAAAAGATG 3359	
DB	804	AAGAAAAAGATG 815	

RESULT 15
 US-10-085-783A-16172
 ; Sequence 16172, Application US/10085783A
 ; Publication No. US20040037841A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ChondroGene Inc.
 ; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
 ; FILE REFERENCE: 4231/2002
 ; CURRENT APPLICATION NUMBER: US/10/085,783A
 ; PRIOR FILING DATE: 2002-02-28
 ; PRIOR APPLICATION NUMBER: US 60/305,340
 ; PRIOR FILING DATE: 2001-07-13
 ; PRIOR APPLICATION NUMBER: US 60/275,017
 ; PRIOR FILING DATE: 2001-03-12
 ; PRIOR APPLICATION NUMBER: US 60/271,955
 ; PRIOR FILING DATE: 2001-02-28
 ; NUMBER OF SEQ ID NOS: 58994
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 16172
 ; LENGTH: 541
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-085-783A-16172

Query Match		13.5%;	Score 499;	DB 12;	Length 541;
Best Local Similarity		96.2%;	Pred. No. 4.1e-117;	Mismatches 511;	Conservative 0;
				Indels 20;	Gaps 0;

QY	2848	AAGCTCAAGACGACGAGCTCAGTGTAGAGGCTAATTCGCTCTGTTGTATTTA	2907
DB	1	AAGCTCAAGACGACGAGCTCAGTGTAGAGGCTAATTCGCTCTGTTGTATTTA	60
QY	2908	TGTTGATTTACTAAATTCGGTTTCATTTATTTTCAATATCCAGTAAACCCATG	2967
DB	61	TGTTGATTTACTAAATTCGGTTTCATTTATTTTCAATATCCAGTAAACCCATG	120
QY	2968	TATATATCATATATTTAATAATCACAGTCTAGAGATGTTTCATGTTAAAGTACTGCT	3027
DB	121	TATATATCATATATTTAATAATCACAGTCTAGAGATGTTTCATGTTAAAGTACTGCT	180
QY	3028	TTGACACGAGCGCTGTTTCTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGAT	3087
DB	181	TTGACACGAGCGCTGTTTCTAAAGAAACCCATGCTGTGAAATAGAGACTTTTCTACTGAT	240
QY	3088	CATCATAACTCTGATCTGAGCAGTGATACCAACCAACATCTGAAGTCAACAGAGATCCA	3147
DB	241	CATCATAACTCTGATCTGAGCAGTGATACCAACCAACATCTGAAGTCAACAGAGATCCA	300
QY	3148	AGTTTAAATTCGCTGCGGAATGTCGAGTATCTAGAAATGAAACCGTAGTTTGTGTT	3207
DB	301	AGTTTAAATTCGCTGCGGAATGTCGAGTATCTAGAAATGAAACCGTAGTTTGTGTT	360
QY	3208	TTTTTAAATACAGAGTCTATGTTTCTGCTACTTTTATTAATTAAGCATGGAAGAAATAT	3267
DB	361	TTTTTAAATACAGAGTCTATGTTTCTGCTACTTTTATTAATTAAGCATGGAAGAAATAT	420
QY	3268	CTTAGTAGGCAATTTGTAACACTTTTGAAGTAAACCCATTCAGATTTGAAATCTCGGA	3327
DB	421	CTTAGTAGGCAATTTGTAACACTTTTGAAGTAAACCCATTCAGATTTGAAATCTCGGA	480

Qy 3328 TAATGGTTGCTTTAAAAAAGAAAAAGATGCGGCGACAGTTATTCTTTT 3378
|||
Db 481 TAATGGTTGCTTTAAAAAAGAAAAAGATGCTGTTAAGGTATTACTTTT 531
|||

Search completed: March 12, 2004, 06:15:01
Job time : 927.748 secs